

JOURNAL  
COMMENT

Unpublished (2003)  
Contact: Tosser-Klopp G  
Genétique Animale  
Institut National de la Recherche Agronomique  
Chemin de Borde-Rouge - Auzeville BP27, 31326 Castanet-Tolosan  
cedex, FRANCE  
Tel: 33 (0) 5.61.28.51.14  
Fax: 33 (0) 5.61.28.53.08  
Email: tosser@toulouse.inra.fr  
Clone distribution: AGENAE Resource centre: Francois PIUMI.  
Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et Etude du  
genome (LREG), Domaine de Vilvert, 78352, Jouy-en-Josas cedex,  
FRANCE, +33 (0) 1.34.65.28.02, +33 (0) 1.34.65.22.73  
Sequence cleaned of vector, adaptor and repetitions. Contact us  
at [sigenasupport@jouy.inra.fr](mailto:sigenasupport@jouy.inra.fr) to obtain the chromatogram of this  
sequence.

FEATURES  
source

29 row: D Corium: 13.  
Location/Qualifiers  
1. .530  
/organism="Sus scrofa"  
/mol\_type="mRNA"  
/db\_xref="taxon:9823"  
/clone="scac00291.b.13"  
/tissue type="mixed"  
/clone lib="Sus Scrofa library (scac)"  
/notes="Vector: p773D-pac vector; tissues: adipose tissue,  
brain, kidney, liver, muscle, ovary, testis, heart,  
hypothalamus, pancreas, skin, spleen, thymus, placenta,  
pituitary gland, seminal vesicle, small intestine,  
uterus, adrenals, bulbo urethral gland, cerebral trunk,  
epididymis, female gonad, gall-bladder, hippocampus,  
large intestine, male gonad, melanocytes, stomach, udder"

## ORIGIN

Alignment Scores:	4.92e-43	Length:	530
Pred. No.:	66.00	Matches:	99
Score:	66.00	Conservative:	0
Percent Similarity:	98.02%	Mismatches:	1
Best Local Similarity:	98.02%	Indels:	2
Query Match:	4.62%	Gaps:	0
DR:	13		

US-09-787-097-12 (1-1429) x BX669805 (1-530)

Qy	146	ThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGlyGln	165
Db	429	ACAGATGGACCTGGAAATTATAAAATATAAAACGAAGTGCATGGGTCTATTGAAGGACAG	370
Qy	166	ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis	185
Db	369	CCAAATAGATAAATCAGACTTCGTTTCAATCATTTTGTACACAGTGTAGTTGGGACCAT	310
Qy	186	LeuTyrValTyrAspGlyAsnSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeu	205
Db	309	TTATATGTTTATGATGGGGACTCAATTTATGCACCTCTAGTCTGCTTTAGCGCCCT	250
Qy	206	IleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyr	225
Db	249	ATCGTTCCTGAGAGAGACAGCAACGAGACTGTCCACAGAGTCTGTTGCCACATCAGGTTA	191
Qy	225	AlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyr	245
Db	190	TGCCCCGTCTGCATTTTTTTTAGCGATGCTGCTTAAATTTGACTGGATTTAAATATCACTTA	131
Qy	245	r 245	
Db	130	C 130	

RESULT 75

Accession	LOCUS	BU506373	964 bp	linear	EST 12-SEP-2002
DEFINITION	AGNCOURT_10015485 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494441				
	5', mRNA sequence.				

ACCESSION	BU506373
VERSION	BU506373.1
KEYWORDS	GI:22812606 EST.
SOURCE	Mus musculus (house mouse)
ORGANISM	Mus musculus

BUS\_musculata  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus  
 1 (bases 1 to 964)  
 NIH-MGC <http://mgc.nci.nih.gov/>.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 COMMENT

Email: cgabbs-r@mail.nih.gov  
Tissue procurement: The Cepko Laboratory  
DNA Library Preparation: Life Technologies  
DNA Library Arrayed by: The I.M.A.G.E.  
DNA Sequencing by: Agencourt Bioscience  
Clone distribution: WGC clone distributor  
found through the I.M.A.G.E. Consortium  
<http://image.llnl.gov>  
Plate: LLN14050 row: k column: 18  
High quality sequence stop: 592.

FEATURES source

```
1. .964
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6494441"
/sequence_time="taxon:10090"
```

```

/tissue_type="retina"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NH10 MC9 94"
/notes="Organ: eye; Vector: pCMV-SPORE6; Site1: Nocl1;
Site 2: Sal1; Cloned unidirectionally; oligo-dT primed.
Average insert size 3.3 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."

```

## ORIGIN

Alignment Scores:	5.7e-42	Length:	964
Pred. No.:	65.00	Matches:	78
Score:	98.73%	Conservative:	0
Percent Similarity:	98.73%	Mismatches:	1
Best Local Similarity:	98.73%	Indels:	0
Query Match:	4.55%	Gaps:	0
DR:	13		

US-09-787-097-12 (1-1429) x BU506373 (1-964)

Qy	842	MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgGly	861
Db	529	ATGCAATCATCTCAAAGTATGTCCAAAGCTCACCTGACTCCATGGGTGGTCTTCGGAG	588
Qy	862	IleAsnValSerTyrTrpCysTrpGluaspMetSerProPheThrAsnSerLeuLeuGln	881
Db	589	ATCAATGTGCTTACTGGTGTGGAGGATATGTCCATTCACAATAGTTTGTGCGAG	648
Qy	882	TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr	901
Db	649	TGGATGCCATCTGACCCACGTGATGCTGGCTCTGTGGGATCTGTGTCAGAGCCTAGTACT	708
Qy	902	Arg-GlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGlu	919
Db	709	CGGGGGATTAAAGGCTGCAACCTCATCAACCTCTCAATGGCAGCGCTGTGAA	763

Search completed: March 1, 2004, 12:54:55  
Job time : 6985 secs

**This Page Blank (uspto)**



IntelliGenetics

FastDB - Fast Pairwise Comparison of Sequences  
Release 5.4

Results file us-09-787-097-12.res made by jdelaval on Sat 17 Jul 104 6:56:22-PST.

Query sequence being compared:US-09-787-097-12 (1-1429)  
 Number of sequences searched: 4  
 Number of scores above cutoff: 3

Results of the initial comparison of US-09-787-097-12 (1-1429) with:  
File : segs-2-10-18.pep

NUMBER OF SEQUENCES

[illegible]

## PARAMETERS

Similarity matrix	Unitary	K-tuple	2
Mismatch penalty	1	Joining penalty	20
Gap penalty	1.00	Window size	32
Gap size penalty	0.05		
Cutoff score	0		
Randomization group	0		

SEARCH STATISTICS

Scores:	Mean	Median	Standard Deviation
	1248	1268	88.05

```

Times:
CPU 00:00:00.00
Total Elapsed 00:00:00.00

```

Number of residues:	3825
Number of sequences searched:	4
Number of scores above cutoff:	3

The scores below are sorted by initial score.

The scores below are sorted by initial score. Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

Sequence Name	Description	Length	Score	Init. Opt.	Sig. Frame
1. US-09-787-097-10	Sequence 10, Application	1355	1325	1331	0.87
2. US-09-787-097-18	Sequence 18, Application	1272	1267	1268	0.22
3. US-09-787-097-2	Sequence 2, Application US	1198	1152	1159	-1.09
	***1 standard deviation below mean***				0
1. US-09-787-097-12	(1-1429)				
US-09-787-097-10	Sequence 10, Application US/09787097				
Sequence 10, Application US/09787097					
GENERAL INFORMATION:					
APPLICANT: Duke-Cohan, Jonathan S.					
APPLICANT: Schlusserman, Stuart F.					
TITLE OF INVENTION: REGULATION OF IMMUNE RESPONSES BY ATTRACTIN					
FILE REFERENCE: 00530-089002					
CURRENT APPLICATION NUMBER: US/09/787,097					
CURRENT FILING DATE: 2001-03-13					
PRIOR APPLICATION NUMBER: PCT/US99/20948					
PRIOR FILING DATE: 1999-09-14					
PRIOR APPLICATION NUMBER: US 60/100,137					
PRIOR FILING DATE: 1998-09-14					
NUMBER OF SEQ ID NOS: 25					
SOFTWARE: FastSeq for Windows Version 3.0					
SEQ ID NO 10					
LENGTH: 1355					
TYPE: PRT					
ORGANISM: Homo sapiens					





THNDTSMGHGAKCFSSDDMAXDIACDRWSVLPRDLDHHDVNRFGHSAVLHNSTWYVFGFNSLLLSILVFT  
THNDTSMGHGAKCFSSDDMAXDIACDRWSVLPRDLDHHDVNRFGHSAVLHNSTWYVFGFNSLLLSILVFT  
500 510 520 530 540 550 560 570  
650 660 670 680 690 700 710  
SEQCDARHSEAAACLAGP3PGIRCVNMTGSSQCIWALATDEQEELKSECFSKRTLDHRCDDQHTDCYSTAN  
SEQCDARHSEAAACLAGP3PGIRCVNMTGSSQCIWALATDEQEELKSECFSKRTLDHRCDDQHTDCYSTAN  
580 590 600 610 620 630 640  
TNDCHWCHDHCVPNRHSCSEGI5IFRYENCPCNDPMYCNKKTSCRSALDQNCQWEPNRQECIALPENIC  
TNDCHWCHDHCVPNRHSCSEGI5IFRYENCPCNDPMYCNKKTSCRSALDQNCQWEPNRQECIALPENIC  
650 660 670 680 690 700 710  
720 730 740 750 760 770 780  
TNDCHWCHDHCVPNRHSCSEGI5IFRYENCPCNDPMYCNKKTSCRSALDQNCQWEPNRQECIALPENIC  
TNDCHWCHDHCVPNRHSCSEGI5IFRYENCPCNDPMYCNKKTSCRSALDQNCQWEPNRQECIALPENIC  
720 730 740 750 760 770 780  
GIGWHLVGN5CLKITAKENTDNAKLCFNRHNALLASLTQKKVEFVLKQLRINQSSQSKLTLTPWVGLR  
GIGWHLVGN5CLKITAKENTDNAKLCFNRHNALLASLTQKKVEFVLKQLRINQSSQSKLTLTPWVGLR  
720 730 740 750 760 770 780  
KINVSYWCEWEDMSPPTNSLLOMPSEPSDAGFCGILSEPTSGILKAATCINPLNGSVYCERPAHNSAKQCRTP  
KINVSYWCEWEDMSPPTNSLLOMPSEPSDAGFCGILSEPTSGILKAATCINPLNGSVYCERPAHNSAKQCRTP  
790 800 810 820 830 840 850 860  
GIGWHLVGN5CLKITAKENTDNAKLCFNRHNALLASLTQKKVEFVLKQLRINQSSQSKLTLTPWVGLR  
GIGWHLVGN5CLKITAKENTDNAKLCFNRHNALLASLTQKKVEFVLKQLRINQSSQSKLTLTPWVGLR  
720 730 740 750 760 770 780  
KINVSYWCEWEDMSPPTNSLLOMPSEPSDAGFCGILSEPTSGILKAATCINPLNGSVYCERPAHNSAKQCRTP  
KINVSYWCEWEDMSPPTNSLLOMPSEPSDAGFCGILSEPTSGILKAATCINPLNGSVYCERPAHNSAKQCRTP  
790 800 810 820 830 840 850 860  
CALRTACGDCTSGSSECCWCSNMKQCVDSNAYVAFPPGQCMWYTWSTCPPENCSCYCTCSHCLCEQPGCGW  
CALRTACGDCTSGSSECCWCSNMKQCVDSNAYVAFPPGQCMWYTWSTCPPENCSCYCTCSHCLCEQPGCGW  
860 870 880 890 900 910 920 930  
CTDPSNTGKGKICIEG5YKGPVMP5QAPTGNFYQPOLNNSMCLEDSDRYNWSFIHCPACQCNHSGKCI NOSI  
CTDPSNTGKGKICIEG5YKGPVMP5QAPTGNFYQPOLNNSMCLEDSDRYNWSFIHCPACQCNHSGKCI NOSI  
940 950 960 970 980 990 1000  
CEKCNLTGKHCECISG5YKGPVMP5QAPTGNFYQPOLNNSMCLEDSDRYNWSFIHCPACQCNHSGKCI NOSI  
CEKCNLTGKHCECISG5YKGPVMP5QAPTGNFYQPOLNNSMCLEDSDRYNWSFIHCPACQCNHSGKCI NOSI  
1010 1020 1030 1040 1050 1060 1070  
CEKCNLTGKHCECISG5YKGPVMP5QAPTGNFYQPOLNNSMCLEDSDRYNWSFIHCPACQCNHSGKCI NOSI  
CEKCNLTGKHCECISG5YKGPVMP5QAPTGNFYQPOLNNSMCLEDSDRYNWSFIHCPACQCNHSGKCI NOSI  
940 950 960 970 980 990 1000  
NPLRGTCYTLIDYQFTFSLQEDDRIYTAINFVATPDEQNRDLDMFINASKNFNLITWAASFSAGTOAG  
NPLRGTCYTLIDYQFTFSLQEDDRIYTAINFVATPDEQNRDLDMFINASKNFNLITWAASFSAGTOAG  
1080 1090 1100 1110 1120 1130 1140  
NPLRGTCYTLIDYQFTFSLQEDDRIYTAINFVATPDEQNRDLDMFINASKNFNLITWAASFSAGTOAG  
NPLRGTCYTLIDYQFTFSLQEDDRIYTAINFVATPDEQNRDLDMFINASKNFNLITWAASFSAGTOAG  
1080 1090 1100 1110 1120 1130 1140  
EEMPVVSNTINIKYKDSN5KXKPFNRHNPITFFVYSNFTWPIKIQIAPFSQHSNFMDLVQFFVTFPFCPLS  
EEMPVVSNTINIKYKDSN5KXKPFNRHNPITFFVYSNFTWPIKIQIAPFSQHSNFMDLVQFFVTFPFCPLS  
1150 1160 1170 1180 1190 1200 1210 1220  
LLLVAAVWVKIKQSCWASRRREQLREMQWASRRFPASVUNVALETDEPPDLIGGSIKTVPKFIALEPCFEN  
LLLVAAVWVKIKQSCWASRRREQLREMQWASRRFPASVUNVALETDEPPDLIGGSIKTVPKFIALEPCFEN  
1300 1310 1320 1330 1340 1350 1360  
KAAVLSVFLRLRGLGIPPGQGLAVASALVDISQMPFIVYKESGAVNRKQPPAQFGTCI  
KAAVLSVFLRLRGLGIPPGQGLAVASALVDISQMPFIVYKESGAVNRKQPPAQFGTCI  
1370 1380 1390 1400 1410 1420

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:07:54 ; Search time 9050.88 Seconds  
(without alignments)  
17225.392 Million cell updates/sec

Title: US-09-787-097-1

Perfect score: 3597  
Sequence: 1 atgggtggcgcagcggcggc.....agggtgcaactgacaatga 3597

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:

1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3597	100.0	3597	9 AF034957	AF034957 Homo sapi
2	3494.4	97.1	4072	6 AR164817	AR164817 Sequence
3	3487	96.9	3819	9 AF166861	AF166861 Homo sapi
4	3476.8	96.7	4589	6 AR164816	AR164816 Sequence
5	3100.8	86.2	4493	4 AF531101	AF531101 Bos tauri
6	2948.8	82.0	6083	10 AF116897	AF116897 Mus muscu
7	2947.8	82.0	4284	10 AB062913	AB062913 Mesocric
8	2913.6	81.0	4313	10 AF119821	AF119821 Mus muscu
9	2909	80.9	8739	10 AB038387	AB038387 Rattus no
10	2907.2	80.8	4500	10 AB038388	AB038388 Rattus no
11	2845.2	79.1	8827	6 AR164807	AR164807 Sequence
12	2255	62.7	5683	10 AB093245	AB093245 Mus muscu
13	2028.8	56.4	2625	6 AR164818	AR164818 Sequence
14	1428.8	39.7	2419	6 AR164813	AR164813 Sequence
15	1364.4	37.9	6370	6 AR164815	AR164815 Sequence
16	1208.4	33.6	2578	9 AK000356	AK000356 Homo sapi
17	1203.2	33.5	4140	9 AF442317	AF442317 Homo sapi
18	1129.4	31.4	3609	6 AX503776	AX503776 Homo sapi
19	879.8	24.5	7302	9 AB011106	AB011106 Homo sapi
20	872.4	24.3	5632	9 AB011120	AB011120 Homo sapi
21	829.2	23.1	5002	10 BC027764	BC027764 Mus muscu
22	817.8	22.7	1012	6 E26747	E26747 Structure a
23	543	15.1	543	6 AX778545	AX778545 Sequence
24	463	12.9	1051	6 AR164814	AR164814 Sequence
25	435.4	12.1	4997	10 BC050020	BC050020 Mus muscu
26	418.4	11.6	1818	9 BC047716	BC047716 Homo sapi
27	340	9.5	3297	10 BC030872	BC030872 Mus muscu
28	253.8	7.1	982	6 AX364943	AX364943 Sequence
29	250.8	7.0	2018	9 H009384308	AF218896 Homo sapi
30	250.8	7.0	104907	9 HSDJ741H3	AL132773 Human DNA
31	250.8	7.0	155869	2 AC023420	AC023420 Homo sapi
32	250.8	7.0	167357	2 AC055771	AC055771 Homo sapi
33	250.8	7.0	206647	9 AP002898	AP002898 Homo sapi
34	250.8	6.8	246	6 AR273223	AR273223 Sequence
35	246	6.8	246	6 AR276804	AR276804 Sequence
36	246	6.8	246	6 AR407079	AR407079 Sequence
37	246	6.8	246	6 AX368256	AX368256 Sequence
38	234.4	6.5	3257	9 H009384815	AF218903 Homo sapi
39	234.4	6.5	118540	2 AC013324	AC013324 Homo sapi
40	234.4	6.5	207433	6 AX326809	AX326809 Sequence
41	234.4	6.2	3589	9 H009384811	AF218899 Homo sapi
42	224.6	6.2	1728	9 H009384814	AF218902 Homo sapi
43	222.6	6.0	237	4 AR273610	AF273610 Sus scrof
44	214.6	5.9	1232	9 H009384805	AF218893 Homo sapi
45	214				

## ALIGNMENTS

RESULT 1

AF034957

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

AF034957 3597 bp mRNA linear PRI 30-SEP-1998  
Homo sapiens secreted T cell activation protein attractin  
(attractin) mRNA, complete cds.

AF034957.1 GI:3676346

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3597)

Duke-Conan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and

Schlossman,S.F.

TITLE Attractin (DpPT-L), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)  
MEDLINE 98409658  
PUBMED 9736737  
REFERENCE 2 (bases 1 to 3597)  
AUTHORS Duke-Cohan, J.S., Gu, J., Ao, Z., McLaughlin, D.P., Freeman, G.J. and Schlossman, S.F.  
DIRECT SUBMISSION  
JOURNAL Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

FEATURES  
source

Location/Qualifiers  
1..3597  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
1..3597  
/gene="attractin"  
1..3597  
/gene="attractin"  
/note="secreted T cell activation protein; Method: conceptual translation with partial peptide sequencing; DpPT-L"  
/codon\_start=1  
/product="Attractin"  
/protein\_id="AAC61902.1"  
/db\_xref="GI:3676347"

## gene

1..3597

## CDS

1..3597  
/translating="WAAAAAATEARLRRTAATAALAGSGGPHCVNGRCNPGTGC  
VCPAGGEGCQHCGRGFTGSGFVTDGPNKYKTKTTLIEQPNRIMRLRFNH  
FATECSWDHLYVDGDSIYAPLAFAAGSLIPERDNGNEVPEVATGVALHFFSDA  
AYNLGTNITYSDPNCNCGRGECKISNSDVECESEWKGACDIPHTDNCG  
FPRGICNSDVRGSCFCSDWQPGCVVPVNPQSWFTREYNSLPLASHKIVYNG  
NIMVVGYNENSDNWLAYDLASREWLPLNRSUNNVVRYGHSALYKLIYMYG  
GKIDSTGNVNLRFVHFNESWLTLPKAEQYAVVGHSAHIVLNGRVNVLVIFG  
HCPYGISYNDLNTWSILKTDGALVQGGYGHSSVYDHRTHALVHGGYKAFS  
AKYRLADLYRYDVTQMTILKDRFFRYLHTAVISGTMVLFQFGNTDTSMSHG  
AKFSDPMDADIAKRWVLPDSTWMTDLAI PAVLHNSMTYVFGGNSLLSIDI  
LVFTSECDADHAEACLAAGPGIRCVMTWSSQICSWALATDEQBEKLKSCFSKRT  
LDHRCDOHTDVCYCTANTNDCHNDHCVNRHNSBGQISIFRYENCPKDNPMYIC  
NKTSCRSALDONCOWEPENORCIAPENICGICKHLVGNCLIKITAKENYDRAKL  
FCRNHALLASLTQKQVFLQRLIMOSSQSMKLTLPWGLKKNVSWCWEDM  
SPTNSLLQWPSPEADGFCGLSPSTRGLKAATCINPLNGSVCCERPAHNSAQCR  
TPCALRTACDCTGSGSECMWCKQCVDSNAYVAFPPGQCMENYTMSTCPENC  
GYCTSCHLEPQPCGCTDFNTGKGCIEGKYGFVPMFQAFTNFPQPLNSM  
CLEDSTRYNSFIHCPACQNGHSCINQISCEKENLTKGHCETCISGFYDPTNGG  
KCQPKCNHSHLNTNTGKCTTKGVKDECOLCEVENRYQGNPLRGTCTYVTLID  
YQFTSLSQSDRDYRTALNFVATPDQNRDLDMFINASKNFNLNTWASFSAGTQAG  
BEMPVSKTNIEKYDSFNSKEDFRNHPNIFVYVSNFTWPIKIQVTEQ"

## ORIGIN

Query Match 100.0%; Score 3597; DB 9; Length 3597;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGTGGCGGAGCGGCGCAATGAGGCAAGCTGAGGAGGAGGAGCGGCGGCGGCA 60  
Db 1 ATGGTGGCGGAGCGGCGCAATGAGGCAAGCTGAGGAGGAGGAGCGGCGGCGGCA 60  
QY 61 GCCTCGCGGCGGAGGAGCGGCGGCGCACTGTGTCAACGGCGGTGCTGCAACCTGGC 120  
Db 61 GCCTCGCGGCGGAGGAGCGGCGGCGCACTGTGTCAACGGCGGTGCTGCAACCTGGC 120  
QY 121 ACCGCGCAGTGTCTGTGCGCGCGGCTGGGTGGGCGAGCAATGCCAGCACTGGCGGGCG 180  
Db 121 ACCGCGCAGTGTCTGTGCGCGCGGCTGGGTGGGCGAGCAATGCCAGCACTGGCGGGCG 180  
QY 181 CGCTTCAGACTAACTGCACTCTTCTGGGTTGTGACAGATGGACCTCGAAATTTATAATAC 240  
Db 181 CGCTTCAGACTAACTGCACTCTTCTGGGTTGTGACAGATGGACCTCGAAATTTATAATAC 240  
QY 241 AAAACGAAGTGACGTTGGCTCAATGAGGACAGCCAAATAGAAATATGAGATCTCGTTTC 300

Db 241 AAAACGAAGTGACGTTGGCTCAATGAGGACAGCCAAATAGAAATATGAGATCTCGTTTC 300  
QY 301 AATCATTTTGTCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATT 360  
Db 301 AATCATTTTGTCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATT 360  
QY 361 ATGCAACCGCTAGTTGCTGCAATTTAGTGGCCTCATTTGTTCTGAGAGAGATGGCAATGAG 420  
Db 361 ATGCAACCGCTAGTTGCTGCAATTTAGTGGCCTCATTTGTTCTGAGAGAGATGGCAATGAG 420  
QY 421 ACTGTCCCTGAGGTTGTTGCCACATTCAGGTTATGCTTGTGCTGCAATTTTATGATGCT 480  
Db 421 ACTGTCCCTGAGGTTGTTGCCACATTCAGGTTATGCTTGTGCTGCAATTTTATGATGCT 480  
QY 481 GCTTATAATTTGACTGGATTTAATAATTACTTACAGTTTGTGATATGTTGCCAAATAACTGC 540  
Db 481 GCTTATAATTTGACTGGATTTAATAATTACTTACAGTTTGTGATATGTTGCCAAATAACTGC 540  
QY 541 TCAGGCGGAGGAGTGTAAAGTCAATGATCAGTAATAGCAGGATCTCTGTAATGTAATGCT 600  
Db 541 TCAGGCGGAGGAGTGTAAAGTCAATGATCAGTAATAGCAGGATCTCTGTAATGTAATGCT 600  
QY 601 GAAAACTCGAAAGGTGAAGCATGTGCAATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660  
Db 601 GAAAACTCGAAAGGTGAAGCATGTGCAATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660  
QY 661 CATCGAGGATCTGCAATTTCAAGTGTAGAGTGTAGAGATGCTCTCTCTCAGACTGGCAG 720  
Db 661 CATCGAGGATCTGCAATTTCAAGTGTAGAGTGTAGAGATGCTCTCTCTCAGACTGGCAG 720  
QY 721 GGTCTCGATGTTCAAGTTCCTGTACAGCAACCACTTTTGACTCTGAGAGGAAATAT 780  
Db 721 GGTCTCGATGTTCAAGTTCCTGTACAGCAACCACTTTTGACTCTGAGAGGAAATAT 780  
QY 781 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGTCAATGGAACATTAATGCTG 840  
Db 781 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGTCAATGGAACATTAATGCTG 840  
QY 841 GTTGTGGAGATATATGTTCAACCACTCAGATTATAACATGTTCTAGAGTATGACCTT 900  
Db 841 GTTGTGGAGATATATGTTCAACCACTCAGATTATAACATGTTCTAGAGTATGACCTT 900  
QY 901 GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAAATGTTGTTAGATATGCT 960  
Db 901 GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAAATGTTGTTAGATATGCT 960  
QY 961 CATTTTGGCAATATACAGGATAAAATTTTACATGTTATGAGGAAAAATGATTCAACT 1020  
Db 961 CATTTTGGCAATATACAGGATAAAATTTTACATGTTATGAGGAAAAATGATTCAACT 1020  
QY 1021 GGGAAATGCAACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCAATGGGTGTTG 1080  
Db 1021 GGGAAATGCAACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCAATGGGTGTTG 1080  
QY 1081 ACCCTTAAGCAAGGAGGAGTATGAGTGGTGGGCACTCTGCAACATTTGTTACTG 1140  
Db 1081 ACCCTTAAGCAAGGAGGAGTATGAGTGGTGGGCACTCTGCAACATTTGTTACTG 1140  
QY 1141 AAGAAATGGCGAGTGGTCACTGCTGCTCATCTTTGGTCACTGCCCTCTCTATGATATA 1200  
Db 1141 AAGAAATGGCGAGTGGTCACTGCTGCTCATCTTTGGTCACTGCCCTCTCTATGATATA 1200  
QY 1201 AGCAATGTGAGGAAATATGATTTGGATTAAGAACACATGGAGTATATTACACCCAGGGT 1260  
Db 1201 AGCAATGTGAGGAAATATGATTTGGATTAAGAACACATGGAGTATATTACACCCAGGGT 1260  
QY 1261 GCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTTTTACACCATAGGACCGGCGCTA 1320  
Db 1261 GCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTTTTACACCATAGGACCGGCGCTA 1320  
QY 1321 TACCTTCATGCTGGCTCAAGAGGCTTTTCAAGTGCCTTCAAGTGCCTTCAAGTGCCTT 1380





QY 3541 GTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAACATGA 3597  
 Db 3541 GTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAACATGA 3597

RESULT 2  
 ARI64817  
 LOCUS ARI64817 4072 bp DNA linear PAT 17-OCT-2001  
 DEFINITION Sequence 16 from patent US 6274339.  
 ACCESSION ARI64817  
 VERSION ARI64817.1 GI:16238027  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 UNCLASSIFIED.  
 REFERENCE 1 (bases 1 to 4072)  
 AUTHORS Moore, K. and Nagle, D. Lynn.  
 TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity  
 JOURNAL Patent: US 6274339-A 16 14-AUG-2001;  
 FEATURES Location/Qualifiers  
 1..4072  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN

Query Match 97.1%; Score 3494.4; DB 6; Length 4072;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 3541; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 15 GCGCGCAACTGAGCAAGGCTGAGAGGAGGAGCGGCGGCGAGCGCTCGCGGCGAG 74  
 Db 234 GGAGGCGAGCGCGCTCGGCTGGCGGCGGCGGTCTCGGCTCGCGCGAGCGGCGCA 293

QY 75 GAGCGGCGGCGCGCACTGTGTCAACGCGGCTGCGTCAACCTGGCAACCGGCGAGCGGT 134  
 Db 294 GGAATGTGACCGGCGGCTGTCAACGCGGCTGCGTCAACCTGGCAACCGGCGAGCGGT 353

QY 135 CTGCGCGCGCGCTGGGCTGGCGGAGCAATCCAGCACTGCGGCGGCGGCTTCAGACTAAC 194  
 Db 354 CTGCGCGCGCGCTGGGCTGGCGGAGCAATCCAGCACTGCGGCGGCGGCTTCAGACTAAC 413

QY 195 TGGATCTCTGGGTTGTGACAGATGCACTGGAAATTAATAACAAACGAAGTGCA 254  
 Db 414 TGGATCTCTGGGTTGTGACAGATGCACTGGAAATTAATAACAAACGAAGTGCA 473

QY 255 GTGGCTCATTTGAAGGAGCGCAATAGAAATAGAGACTCGTTTCAATCTTTGCTAC 314  
 Db 474 GTGGCTCATTTGAAGGAGCGCAATAGAAATAGAGACTCGTTTCAATCTTTGCTAC 533

QY 315 AGAGTGTAGTTGGGACCAATTTATATGTTATGATGGGAGTCAATTTATGCAACCGTACT 374  
 Db 534 AGAGTGTAGTTGGGACCAATTTATATGTTATGATGGGAGTCAATTTATGCAACCGTACT 593

QY 375 TGCTGCAATTTAGTGGCTCTATGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434  
 Db 594 TGCTGCAATTTAGTGGCTCTATGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653

QY 435 TGTGGCACATCAGGTTATGCTGTGCAATTTTATGATGCTGTGCTTATAATTTGAC 494  
 Db 654 TGTGGCACATCAGGTTATGCTGTGCAATTTTATGATGCTGTGCTTATAATTTGAC 713

QY 495 TGAATTTAATATTAATTAAGTTTGGATGATGCTGCAAAATACTGCTAGGCGGAGAGA 554  
 Db 714 TGAATTTAATATTAATTAAGTTTGGATGATGCTGCAAAATACTGCTAGGCGGAGAGA 773

QY 555 GTGTAAGATCAGTAATAGCAGCGATCTGTGTAATGTGAATGTTCTGAAAACCTGGAAGG 614  
 Db 774 GTGTAAGATCAGTAATAGCAGCGATCTGTGTAATGTGAATGTTCTGAAAACCTGGAAGG 833

QY 615 TGAAGCATGTGACATTTCTCTACCTGTATACAGCAACTGTGTGTTTCTCTCATCGAGGATCTG 674  
 Db 834 TGAAGCATGTGACATTTCTCTACCTGTATACAGCAACTGTGTGTTTCTCTCATCGAGGATCTG 893

QY 675 CAATTCAAGTGTATGTCAAGAGATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGATGTTTC 734  
 Db 894 CAATTCAAGTGTATGTCAAGAGATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGATGTTTC 953

QY 735 AGTTCTGTACCAAGTAAACCAAGTCAATTTTGGACTCGAGAGAAATATTCTAACTTAAAGCT 794  
 Db 954 AGTTCTGTACCAAGTAAACCAAGTCAATTTTGGACTCGAGAGAAATATTCTAACTTAAAGCT 1013

QY 795 CCCAGAGCATCTCATAAAGCTGTGCTCAATGGAAACATTTATGTTGGTGTGAGAGATA 854  
 Db 1014 CCCAGAGCATCTCATAAAGCTGTGCTCAATGGAAACATTTATGTTGGTGTGAGAGATA 1073

QY 855 TATGTTCAACCACTCAGATTAATAACATGTTCTAGCGTATGACCTTGTCTTAGGGAGTG 914  
 Db 1074 TATGTTCAACCACTCAGATTAATAACATGTTCTAGCGTATGACCTTGTCTTAGGGAGTG 1133

QY 915 GCTTCCACTAAACCGTCTGTGAACAATGTTGTTAGATATGTTCTTCTTTGGCATTT 974  
 Db 1134 GCTTCCACTAAACCGTCTGTGAACAATGTTGTTAGATATGTTCTTCTTTGGCATTT 1193

QY 975 ATACAAGGATAAATTTACATGTTATGAGGAGAAATTTGATTCAACTGGGAATGTGACAA 1034  
 Db 1194 ATACAAGGATAAATTTACATGTTATGAGGAGAAATTTGATTCAACTGGGAATGTGACAA 1253

QY 1035 TGAGTTGAGAGTTTTCATCTCATATAGTCAATGAGTGTGTTGACCTTACCCCTAAGGCAAA 1094  
 Db 1254 TGAGTTGAGAGTTTTCATCTCATATAGTCAATGAGTGTGTTGACCTTACCCCTAAGGCAAA 1313

QY 1095 GGAGCAGTATCAGTGGTGGGCACTCTGCACACATTTGTACACTGAAGAATGSCCGAGT 1154  
 Db 1314 GGAGCAGTATCAGTGGTGGGCACTCTGCACACATTTGTACACTGAAGAATGSCCGAGT 1373

QY 1155 GGTCAATGTGTCTATCTTGGTCACTGCCCTCTCTATGATGATATATAAGCAATGTGCAGGA 1214  
 Db 1374 GGTCAATGTGTCTATCTTGGTCACTGCCCTCTCTATGATGATATATAAGCAATGTGCAGGA 1433

QY 1215 ATATGATTTGGATTAAGAACACATCGAGTATATTACACACCCAGGGTCCCTTGTGCAAGG 1274  
 Db 1434 ATATGATTTGGATTAAGAACACATCGAGTATATTACACACCCAGGGTCCCTTGTGCAAGG 1493

QY 1275 GGGTTACGGCCATAGCAGTGTTCACGACCATAGGACCCAGGGCCCTATACGTTTCATGTTGG 1334  
 Db 1494 GGGTTACGGCCATAGCAGTGTTCACGACCATAGGACCCAGGGCCCTATACGTTTCATGTTGG 1553

QY 1335 CTACAGGCTTTCACTGGCAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394  
 Db 1554 CTACAGGCTTTCACTGGCAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613

QY 1395 GGATACCCAGATGTGGACCAATTTCTTAAGGACAGCCGATTTTCCGTTTACITGTCACACAGC 1454  
 Db 1614 GGATACCCAGATGTGGACCAATTTCTTAAGGACAGCCGATTTTCCGTTTACITGTCACACAGC 1673

QY 1455 TGTGATAGTGTGAGGAAACCATGCTGTGTTTGGGGGAAAACACACAAATGACACATCTAT 1514  
 Db 1674 TGTGATAGTGTGAGGAAACCATGCTGTGTTTGGGGGAAAACACACAAATGACACATCTAT 1733

QY 1515 GAGCCATGGCGCCCAATGCTTCTCTCAGATTTTCAATGGCCCTATGCAATTCGCTGTGACCG 1574  
 Db 1734 GAGCCATGGCGCCCAATGCTTCTCTCAGATTTTCAATGGCCCTATGCAATTCGCTGTGACCG 1793

QY 1575 CTGCTCAGTGTCTCCAGAGCTGAT-TCCACCAATGATGTCAAAGATTTGGCCATTTCCAG 1633  
 Db 1794 CTGCTCAGTGTCTCCAGAGCTGAT-TCCACCAATGATGTCAAAGATTTGGCCATTT-CAG 1852

QY 1634 CAGTCTTTACACAACAGCACCATGATGTGTTTGGTGTGTTTCAATAGTCTCTCTCTCAGCG 1693  
 Db 1853 CAGTCTTTACACAACAGCACCATGATGTGTTTGGTGTGTTTCAATAGTCTCTCTCTCAGCG 1912

QY 1694 ACATCTGGTATTCACTCTCGAAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTGTTAG 1753  
 Db 1913 ACATCTGGTATTCACTCTCGAAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTGTTAG 1972



QY 1754 CAGCAGACCTGGTATTTCGGTGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCGTGGG 1813  
Db 1753 CAGCAGACCTGGTATTTCGGTGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCGTGGG 2032  
QY 1814 CGCTGGCACTGATGACAGAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTC 1873  
Db 2033 CGCTGGCACTGATGACAGAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTC 2092  
QY 1874 TTGACCATGACAGATGTGACAGCAGACACAGATTGTTTACAGCTGCACAGGCCAACACCAATG 1933  
Db 2093 TTGACCATGACAGATGTGACAGCAGACACAGATTGTTTACAGCTGCACAGGCCAACCAATG 2152  
QY 1934 ACTGCCACTGGTGAATGACATGTGTCCAGGAAACACAGCTGCTCAGAGGCCAGA 1993  
Db 2153 ACTGCCACTGGTGAATGACATGTGTCCAGGAAACACAGCTGCTCAGAGGCCAGA 2212  
QY 1994 TCTCCATTTTGTAGTATGAAATGCCCCAAGGATAACCTATGTACTACTGTAAACAAG 2053  
Db 2213 TCTCCATTTTGTAGTATGAAATGCCCCAAGGATAACCTATGTACTACTGTAAACAAG 2272  
QY 2054 AGACCACTGACAGAGCTGTGCCCTGGAACAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2113  
Db 2273 AGACCACTGACAGAGCTGTGCCCTGGAACAGAACTGCCAGTGGGAGCCCCGGAATCAGG 2332  
QY 2114 AGTGCAATGCTGCCCGGAAATATCTGTGSCATGTGCGCATTTGGTGTGAAACTCAT 2173  
Db 2333 AGTGCAATGCTGCCCGGAAATATCTGTGSCATGTGCGCATTTGGTGTGAAACTCAT 2392  
QY 2174 GTTTGAAATTTACTACTGCCAAGGAGAAATATGACAAATGTCATAATGTTCTGTAGGAAC 2233  
Db 2393 GTTTGAAATTTACTACTGCCAAGGAGAAATATGACAAATGTCATAATGTTCTGTAGGAAC 2452  
QY 2234 ACAATGCCCTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGTCCTTTAAGCAGC 2293  
Db 2453 ACAATGCCCTTTGGCTTCTCTTACACCCAGAGAGGTAGAAATTTGTCCTTTAAGCAGC 2512  
QY 2294 TGGCAATTAATGCACTCATCTCAGAGCATGTCGAAGTCCACCTTAAACCCATGGTGGGCC 2353  
Db 2513 TGGCAATTAATGCACTCATCTCAGAGCATGTCGAAGTCCACCTTAAACCCATGGTGGGCC 2572  
QY 2354 TTGGAAGATCAATGTGTCTCTA CTGTGCTGGGAAGATGTCCCATTTTACAAATAGTT 2413  
Db 2573 TTGGAAGATCAATGTGTCTCTA CTGTGCTGGGAAGATGTCCCATTTTACAAATAGTT 2632  
QY 2414 TACTACAGTGTGTCGCTGTAGCCAGTCATGTCGATTCGTGGAATTTTATCAGAAC 2473  
Db 2633 TACTACAGTGTGTCGCTGTAGCCAGTCATGTCGATTCGTGGAATTTTATCAGAAC 2692  
QY 2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTGTGTCTGTG 2533  
Db 2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTGTGTCTGTG 2752  
QY 2534 AAGGCTGCNAACCAACAGTCTAAGCAGTCCCGACACCATGTGCTTGGAGACAGCAT 2593  
Db 2753 AAGGCTGCNAACCAACAGTCTAAGCAGTCCCGACACCATGTGCTTGGAGACAGCAT 2812  
QY 2594 GTGAGATTTGACAGCGGAGCTCTCAGTGCATGTGTCAGCAGCAACATGAACAGTGTG 2653  
Db 2813 GTGAGATTTGACAGCGGAGCTCTCAGTGCATGTGTCAGCAGCAACATGAACAGTGTG 2872  
QY 2654 TGGACTCCANTGCTGTGCTCTCTTCCCTTTGGCCAGTGTATGGAATGTATACGA 2713  
Db 2873 TGGACTCCANTGCTGTGCTCTCTTCCCTTTGGCCAGTGTATGGAATGTATACGA 2932  
QY 2714 TGACACTGCCCCCTGAAATTTTTCAGCTACTGTACTGTAGTCAATGCTTGGAGC 2773  
Db 2933 TGACACTGCCCCCTGAAATTTTTCAGCTACTGTACTGTAGTCAATGCTTGGAGC 2992  
QY 2774 AACAGGCTGTGGTGTGCTACTGATCCAGCAATACTGCAAGGGAATGATAGAGG 2833  
Db 2993 AACAGGCTGTGGTGTGCTACTGATCCAGCAATACTGCAAGGGAATGATAGAGG 3052  
QY 2834 GTTCTATAAGGACCAAGTGAAGTGCCTTCGCAAGCCCCCTACAGGAAATTTCTATCCAC 2893

Db 3053 GTTCTATAAGGACCAAGTGAAGTGCCTTCGCAAGCCCCCTACAGGAAATTTCTATCCAC 3112  
QY 2894 AGCCCTCTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 2953  
Db 3113 AGCCCTCTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCATTC 3172  
QY 2954 ACTGTCCAGCTTGGCAATGCAACCGCCACAGTAATGTCATCAATCAGAGCATCTGTGAGA 3013  
Db 3173 ACTGTCCAGCTTGGCAATGCAACCGCCACAGTAATGTCATCAATCAGAGCATCTGTGAGA 3232  
QY 3014 AGTGTGGAACCTGCACACAGCAGCAAGCTCGAGAGCTGCATATCTGGCTTCTACGGTG 3073  
Db 3233 AGTGTGGAACCTGCACACAGCAGCAAGCTCGAGAGCTGCATATCTGGCTTCTACGGTG 3292  
QY 3074 ATCCCAACCAATGGAGGAAATGTGAGCCATGCAATGCAATGGGACACGCTCTCTGTGCA 3133  
Db 3293 ATCCCAACCAATGGAGGAAATGTGAGCCATGCAATGCAATGGGACACGCTCTCTGTGCA 3352  
QY 3134 ACACCAACACGGGCAAGTCTTCTGCACCAAGGGGCTCAAGGGGACAGTGGCCAGC 3193  
Db 3353 ACACCAACACGGGCAAGTCTTCTGCACCAAGGGGCTCAAGGGGACAGTGGCCAGC 3412  
QY 3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTACTC 3253  
Db 3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTACTC 3472  
QY 3254 TTCTTATTGACTATCAGTTTCACTTTAGTCTATCCAGGAAAGATGATCGCTATTACACAG 3313  
Db 3473 TTCTTATTGACTATCAGTTTCACTTTAGTCTATCCAGGAAAGATGATCGCTATTACACAG 3532  
QY 3314 CTATCAATTTTGTGGTACTCTCTGACGAAACAAACAGGATTTGACATGTTTCATCAATG 3373  
Db 3533 CTATCAATTTTGTGGTACTCTCTGACGAAACAAACAGGATTTGACATGTTTCATCAATG 3592  
QY 3374 CTTCAAGAAATTTCAACCTCAACATCACTCGGCTGCCAGTTTCTCAGCTGGAAACCCAGG 3433  
Db 3593 CTTCAAGAAATTTCAACCTCAACATCACTCGGCTGCCAGTTTCTCAGCTGGAAACCCAGG 3652  
QY 3434 CTGGAGAGAGATGCTCTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493  
Db 3653 CTGGAGAGAGATGCTCTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3712  
QY 3494 CTAATGAGAAGTTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTA 3553  
Db 3713 CTAATGAGAAGTTTGATTTTGGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTA 3772  
QY 3554 ATTTCACTGCCCCTCAAAAATTCAGGTGCAAACTGAACAATGA 3597  
Db 3773 ATTTCACTGCCCCTCAAAAATTCAGGTGCAAACTGAACAATGA 3816

RESULT 3  
AF106861

LOCUS AF106861 3819 bp mRNA linear PRI 06-JAN-1999

DEFINITION Homo sapiens attractin-2 (ATR2) mRNA, complete cds.

ACCESSION AF106861

VERSION AF106861.1 GI:4093195

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 3819)

AUTHORS Duke-Cohan, J.S., Gu, J., Freeman, G.J. and Schlossman, S.F.

TITLE Cloning of cDNA for attractin-2, identical with that of attractin

except for a GC-rich 222 bp 5' insertion

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 3819)

AUTHORS Duke-Cohan, J.S., Gu, J., Freeman, G.J. and Schlossman, S.F.

TITLE Direct Submission

JOURNAL Submitted (15-NOV-1998) Division of Tumor Immunology, Dana Farber

Cancer Institute, 44 Binney Street, Boston, MA 02115, USA



Db 1770 CATGGCCTATGACATTCCTGTGACCGCTGGTCAAGTGTCCAGAGACCTGATCTCCACCA 1829  
Qy 1607 TGAATGCAACAGATTTGGCCATTCCAGCAGTCTTACACAAACAGCACCAGTATGTTGCG 1666  
Db 1830 TGAATGCAACAGATTTGGCCATT-CAGCAGTCTTACACAAACAGCACCAGTATGTTGCG 1888  
Qy 1667 GTGGTTTCAATAGTCTCCTCCTCAGCGAATCTCTGTATATACCTCGGAAACAGTGTGATG 1726  
Db 1889 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCTGTGATATTCACCTCGGAAACAGTGTGATG 1948  
Qy 1727 CGCATCGGAGTGAAGCCGCTGTTTATAGCAGCAGGACCTGTATTCGTTGTGTGGAAACA 1786  
Db 1949 CGCATCGGAGTGAAGCCGCTGTTTATAGCAGCAGGACCTGTATTCGTTGTGTGGAAACA 2008  
Qy 1787 CAGGGTCTGTCTAGTGTATCTCTGTGGCGCTGGCAACTGATGAACAAAGAAAGAAATTTAA 1846  
Db 2009 CAGGGTCTGTCTAGTGTATCTCTGTGGCGCTGGCAACTGATGAACAAAGAAAGAAATTTAA 2068  
Qy 1847 AATCAGATGTTTTTCCAAAGAACTCTTGAACATGACAGATGTGACCAAGCAACAGATT 1906  
Db 2069 AATCAGATGTTTTTCCAAAGAACTCTTGAACATGACAGATGTGACCAAGCAACAGATT 2128  
Qy 1907 GTTACAGCTGTACAGCCAAACCAACATGACTGGCCACTGGTGAATGACCAATTTGTGCCCA 1966  
Db 2129 GTTACAGCTGTACAGCCAAACCAACATGACTGGCCACTGGTGAATGACCAATTTGTGCCCA 2188  
Qy 1967 GGAACACAGCTCTCAGAAAGCCAGATCTCCAATTTTAGGTATGAGAAATTTGCCCAAG 2026  
Db 2189 GGAACACAGCTCTCAGAAAGCCAGATCTCCAATTTTAGGTATGAGAAATTTGCCCAAG 2248  
Qy 2027 ATAAACCTATGTACTACTGTAACAAAGAACACAGCTGACAGGAGCTGTGCCCTGGACCA 2086  
Db 2249 ATAAACCTATGTACTACTGTAACAAAGAACACAGCTGACAGGAGCTGTGCCCTGGACCA 2308  
Qy 2087 ACTGCCAGTGGAGCCCGGAAATCAGAGTGCATTTGCCCTGCCGAAATATCTGTGGCA 2146  
Db 2309 ACTGCCAGTGGAGCCCGGAAATCAGAGTGCATTTGCCCTGCCGAAATATCTGTGGCA 2368  
Qy 2147 TTGGCTGGCAATTTGGTTGGAATCTCATGTTTGAATTTACTGTGCCAAGGAAATATG 2206  
Db 2369 TTGGCTGGCAATTTGGTTGGAATCTCATGTTTGAATTTACTGTGCCAAGGAAATATG 2428  
Qy 2207 ACAATGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCTTCTCTTACAAACCA 2266  
Db 2429 ACAATGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCTTCTCTTACAAACCA 2488  
Qy 2267 AGAAGTAGAATTTGTTCTTAAAGCAGCTGCAATATGAGTCAATCTCAGAGCATGTCCA 2326  
Db 2489 AGAAGTAGAATTTGTTCTTAAAGCAGCTGCAATATGAGTCAATCTCAGAGCATGTCCA 2548  
Qy 2327 AGCTCAGCTTAAACCCCATGGTGGCTTCGGAAGATCAATGTGTCTTACTGTGTGTGG 2386  
Db 2549 AGCTCAGCTTAAACCCCATGGTGGCTTCGGAAGATCAATGTGTCTTACTGTGTGTGG 2608  
Qy 2387 AAGATATGCCCCATTTTACAAATAGTTTACTACAGTGGATGCGCTGTAGCCCAAGTATG 2446  
Db 2609 AAGATATGCCCCATTTTACAAATAGTTTACTACAGTGGATGCGCTGTAGCCCAAGTATG 2668  
Qy 2447 CTGGATTTCTGGGAATTTTATCAGAACCCAGTACTCGGGAGCTGAAGCTGCAACCTGCA 2506  
Db 2669 CTGGATTTCTGGGAATTTTATCAGAACCCAGTACTCGGGAGCTGAAGCTGCAACCTGCA 2728  
Qy 2507 TCACACCCACTCAATGGTAGTGTCTGTGAAAGGCTGCAAAACCAAGTGTCTTAAAGCAGTGC 2566  
Db 2729 TCACACCCACTCAATGGTAGTGTCTGTGAAAGGCTGCAAAACCAAGTGTCTTAAAGCAGTGC 2788  
Qy 2567 GGACACCATGTGCTTGGAGCAGCATGTGGAGATTTGACACAGCGGAGCTGTAGTGCA 2626  
Db 2789 GGACACCATGTGCTTGGAGCAGCATGTGGAGATTTGACACAGCGGAGCTGTAGTGCA 2848  
Qy 2627 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCTATGTGGCTCTTCCCTT 2686  
Db 2849 TGTGGTGCAGCAACATGAAGCAGTGTGTGGACTCCAATGCTATGTGGCTCTTCCCTT 2908

Qy 2687 TTGCCAGTGTATGGAATGGTATACGATGAGACCTGCCCCCTGAAAAATTTGTTCAGGCT 2746  
Db 2909 TTGCCAGTGTATGGAATGGTATACGATGAGACCTGCCCCCTGAAAAATTTGTTCAGGCT 2968  
Qy 2747 ACTGTACTGTAGTCAATTTGTTGGAGCAACAGGCTGTGGTGTGTACTGATCCACGCA 2806  
Db 2969 ACTGTACTGTAGTCAATTTGTTGGAGCAACAGGCTGTGGTGTGTACTGATCCACGCA 3028  
Qy 2807 ATACTGCAAGAGGAAATGATAGAGGTTCTTATAAGGACCAAGTGAAGATGCTTCGC 2866  
Db 3029 ATACTGCAAGAGGAAATGATAGAGGTTCTTATAAGGACCAAGTGAAGATGCTTCGC 3088  
Qy 2867 AAGCCCCCTACAGGAAATTTCTATCCACAGCCCCCTCAATTCACAGATGTGTCTAGAG 2926  
Db 3089 AAGCCCCCTACAGGAAATTTCTATCCACAGCCCCCTCAATTCACAGATGTGTCTAGAG 3148  
Qy 2927 ACAGCAGATCAAACTGTGTTTCAATTCATCTGTCAGCTTGCCAAATGCAAGGCAACAGTA 2986  
Db 3149 ACAGCAGATCAAACTGTGTTTCAATTCATCTGTCAGCTTGCCAAATGCAAGGCAACAGTA 3208  
Qy 2987 AATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCAAGGCAAGCACTGCG 3046  
Db 3209 AATGATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCAAGGCAAGCACTGCG 3268  
Qy 3047 AGACTGCTATCTGGCTTCTACGGTGATCCCAATGGAGGAAATGTGAGCCATGCA 3106  
Db 3269 AGACTGCTATCTGGCTTCTACGGTGATCCCAATGGAGGAAATGTGAGCCATGCA 3328  
Qy 3107 AGTCAATGCGCAGCGCTCTGTGCAACACACACAGGCAAGTGTCTTGCACCA 3166  
Db 3329 AGTCAATGCGCAGCGCTCTGTGCAACACACACAGGCAAGTGTCTTGCACCA 3388  
Qy 3167 AGGGGCTCAAGGGGAGGAGTGCAGCTATGTGAGGTAGAAAAATCGATACCAAGAAACC 3226  
Db 3389 AGGGGCTCAAGGGGAGGAGTGCAGCTATGTGAGGTAGAAAAATCGATACCAAGAAACC 3448  
Qy 3227 CTCTCAGAGGAACATGTTTATATATCTCTTTATGACTATCAGTTACCTTTTATCTAT 3286  
Db 3449 CTCTCAGAGGAACATGTTTATATATCTCTTTATGACTATCAGTTACCTTTTATCTAT 3508  
Qy 3287 CCCAGGAGATGATGCTTATATACAGCTATCAATTTTGTGCTACTCTCAGCAACAAA 3346  
Db 3509 CCCAGGAGATGATGCTTATATACAGCTATCAATTTTGTGCTACTCTCAGCAACAAA 3568  
Qy 3347 ACAGGAAATTTGGACATGTTTCAATGCTCCAAAGAAATTTCAACCTCAACATCACCTGG 3406  
Db 3569 ACAGGAAATTTGGACATGTTTCAATGCTCCAAAGAAATTTCAACCTCAACATCACCTGG 3628  
Qy 3407 CTGCAGATTTCTCAGCTGGAAACCCAGGCTGAGAGAGATGCTGTTGTTTCAAAAACCA 3466  
Db 3629 CTGCAGATTTCTCAGCTGGAAACCCAGGCTGAGAGAGATGCTGTTGTTTCAAAAACCA 3688  
Qy 3467 ACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAAGTTTGAATTTTCCAAACCA 3526  
Db 3689 ACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAAGTTTGAATTTTCCAAACCA 3748  
Qy 3527 ATATCATCTTTCTTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGGTGCAAA 3586  
Db 3749 ATATCATCTTTCTTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGGTGCAAA 3808  
Qy 3587 CTGAACAATGA 3597  
Db 3809 CTGAACAATGA 3819

## RESULT 4

LOCUS

AR164816

DEFINITION

Sequence 14 from patent US 6274339.

ACCESSION

AR164816

VERSION

AR164816.1

KEYWORDS

GI:16238025

8589 bp

DNA

linear

PAT 17-OCT-2001



```
1934 ACTGCCACTGGTGCAATGACCAATGTGTCTCCACAGGAACCAACAGCTGCTCAGAGGCCAGA 1993
Db ACTGCCACTGGTGCAATGACCAATGTGTCTCCACAGGAACCAACAGCTGCTCAGAGGCCAGA 2212
1994 TCTCCCAATTTTGGTATGAGATTCGCCCAAGGATTAACCTATGTAATCTCTGTAACAAGA 2053
Db TCTCCCAATTTTGGTATGAGATTCGCCCAAGGATTAACCTATGTAATCTCTGTAACAAGA 2272
2054 AGACCAAGCTGCGAGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGCCCGGAATCAGG 2113
Db AGACCAAGCTGCGAGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGCCCGGAATCAGG 2332
2114 AGTGCAATGCTGCGCGAAATATCTGGCAATCTGGCTGGCAATTTGGTTGGAACTCAT 2173
Db AGTGCAATGCTGCGCGAAATATCTGGCAATCTGGCTGGCAATTTGGTTGGAACTCAT 2392
2174 GTTTGAAATTTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGAAAC 2233
Db GTTTGAAATTTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGAAAC 2452
2234 ACAATGCCCTTTGGCTTCTTTACAAACCAAGAGGTAGAAATTTGTCCTTTAGCAGC 2293
Db ACAATGCCCTTTGGCTTCTTTACAAACCAAGAGGTAGAAATTTGTCCTTTAGCAGC 2512
2294 TGGCAATATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGGCC 2353
Db TGGCAATATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGGCC 2572
2354 TTCGGAAGTCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAGTT 2413
Db TTCGGAAGTCAATGTGCTCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAATAGTT 2632
2414 TACTACAGTGGATGCGCTGTCAGCCAGTGCATGCTGGATCTCTGGAATTTATCAGAAC 2473
Db TACTACAGTGGATGCGCTGTCAGCCAGTGCATGCTGGATCTCTGGAATTTATCAGAAC 2692
2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGTG 2533
Db CCAGTACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGTG 2752
2534 AAGAGCTGCAAAACCAAGTGTGAAGCAGTCCGAGACCAATGCTGCTTGGAGACAGCAT 2593
Db AAGAGCTGCAAAACCAAGTGTGAAGCAGTCCGAGACCAATGCTGCTTGGAGACAGCAT 2812
2594 GTGAGATTCACACAGCGGAGCTGTGAGTGCATGTGTGACCAACACATGAAGCAGTGTG 2653
Db GTGAGATTCACACAGCGGAGCTGTGAGTGCATGTGTGACCAACACATGAAGCAGTGTG 2872
2654 TGGACTCCAATGCTATGTCCTTCTTCCCTTTTGGCCAGTGTATGGAATGATATACGA 2713
Db TGGACTCCAATGCTATGTCCTTCTTCCCTTTTGGCCAGTGTATGGAATGATATACGA 2932
2714 TGAGCAGCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTTGGAGC 2773
Db TGAGCAGCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTTGGAGC 2992
2774 AACAGGCTGTGGTGTGCTACTGATCCAGCAATATCTGCAAGGGAATGATAGAGC 2833
Db AACAGGCTGTGGTGTGCTACTGATCCAGCAATATCTGCAAGGGAATGATAGAGC 3052
2834 GTTCCATAAGGACAGTGAAGTGTGTCAGGCTTCCAGGCCCTACAGGAATTTCTATCCAC 2893
Db GTTCCATAAGGACAGTGAAGTGTGTCAGGCTTCCAGGCCCTACAGGAATTTCTATCCAC 3112
2894 AGCCCTGCTCAATTTCCAGATGTGTCTAGAGCAGCAGATCAATCGCTTTTCTATTC 2953
Db AGCCCTGCTCAATTTCCAGATGTGTCTAGAGCAGCAGATCAATCGCTTTTCTATTC 3172
2954 ACTGTCCAGCTGTGCAATGCAACGGCCACAGTAAATGCAATCAATCAAGCATCTGTGAGA 3013
Db ACTGTCCAGCTGTGCAATGCAACGGCCACAGTAAATGCAATCAATCAAGCATCTGTGAGA 3232
3014 AGTGTGAGAACCTGACCAACAGGACAGCAGTCCGAGACCTGCATATCTCGGCTTCTACGGTG 3073
```

```
1000
3233 AGTGTGAACTGACCAACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTG 3292
3074 ATCCACCAATGGAGGAAATGTGAGCATGCAAGTGAATGGCAGCGCTCTCTGTGCA 3133
3293 ATCCACCAATGGAGGAAATGTGAGCATGCAAGTGAATGGCAGCGCTCTCTGTGCA 3352
3334 ACACCAACAGGGCAAGTGTGTCACCAAGGCGCTCAAGGGGAGCAGTGGCAGC 3193
3353 ACACCAACAGGGCAAGTGTGTCACCAAGGCGCTCAAGGGGAGCAGTGGCAGC 3412
3194 TATGTGAGTGAAGAAATGATACCAAGGAACCTCTCAGAGGAACATGTTATTTATCTC 3253
3413 TATGTGAGTGAAGAAATGATACCAAGGAACCTCTCAGAGGAACATGTTATTTATCTC 3472
3254 TTCTATTGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTTATTTACAG 3313
3473 TTTCTATTGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTTATTTACAG 3532
3314 CTATCAATTTTGTGGCTACTCTGACGAACAAACAGGATTTGACATGTTTCATCAATG 3373
3533 CTATCAATTTTGTGGCTACTCTGACGAACAAACAGGATTTGACATGTTTCATCAATG 3592
3374 CTTCAAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGTGGAAACCCAGG 3433
3593 CTTCAAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGTGGAAACCCAGG 3652
3434 CTGGAAGAGATGCTGCTGTTGTTTCAAAACCAACATTAAGGATGACAAAGATAGTTTCT 3493
3653 CTGGAAGAGATGCTGCTGTTGTTTCAAAACCAACATTAAGGATGACAAAGATAGTTTCT 3712
3494 CTAATGAGAACTTTGATTTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTA 3553
3713 CTAATGAGAACTTTGATTTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTA 3772
3554 ATTTCACCTGCGCCCAATCAAAATTCAGGT 3581
3773 ATTTTCACCTGCGCCCAATCAAAATTCAGAT 3800
```

## RESULT 5

AF531101 Bos taurus attractin mRNA, complete cds.

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

AF531101.1 GI:22297307

SDAAVNLTFGNTIYNPDVPCNCSGRGECKI SNSNTVOCESSENKGBACDIPHCN  
NCGPPHGI CNSSDVRGSCSPSEWQPGSCSPVPANOSFWTRREVSDLLKLRASHKAV  
VNGNIMWVGVMFNDSYNNWLYDLASREMLALARSVNSVVRVYGHSLALYQDKIY  
MYGKIDSTGNINELRVFIRHNSWLLSPKAEQTAUVGHSAAHIVTLKNGRVMLV  
IFGHCPLYGYSISVQVEYDLDNTWLSILHTQGLYVQGGHSSVYDRTKALYIHGGYK  
AFSANKYRLADDDIYRDYDQMTWILKDSFRFYRLHTAVIVSGTMLVFGGNTHTDSM  
SHGAKCFSSFMAYDIDCDRWSVLPRLDHLHDVNRFGHSAVLNRYLVFGGNSLL  
SDILVTFSEOECAHOSBAACLAAGPVRGMDTGSQCVSWELAPAEQKI KSECFPSK  
RI FDHBCDOLTDYCSANTNGCQNDHCVPNHSCTEGOLS I FKYDHCPRKPNMY  
YCNKYSRCRALDQOCWEPRNQECIALPENICGIGHWLVGNSCLKIITARETDYNA  
KLSRHNENALSLATOKKVFVKQLRIMQSSQMSKFTLTPWGLKRNINVSICWE  
DMSPFTNSSIOWMPSFSDAGFCGLISEPSTRGLKAATCINPLNGSVCEPRAHSAKQ  
CRTPFCALRTAGBCTSGSSCCSMNMKQCVDSNAYVAFPFQGCMEWTMSSCPPEN  
CSGYCTCSHLEQCGCWCTPSNTGKICI EGSYKGPVMPBSQGTGNSYPOPLINS  
SMCLDSRYNWSIFHPACOCNGHSKINOSICEKENLITGKHCTCISGFVGDPTN  
GGKOPCEONGHASLONTWTKCFCTTKGVKDECOLCEVENYQONPLKTCYVTL  
IDVQFTSLQEDDRYYTAINFVATPDEQNRDLDMFINASKFNINLITWAASFAGTQ  
AGBMPVVSNTIKKEYKDSFNSNEKFPDRHNPNTFFVYVSNFTWPIKI QIAFSQHSN  
MDLVQFVTFPFCFLSLLVAVVMKIKOSCWASRRERQLLREMOQASRPASVNVVA  
LETDEBPDLIGGSIKTPKPIALEPCFNGKAAVLSVFLPRGLGIGIPPGQSGGLAV  
ASALVDISQMPVYKESGAVNRXQPPQPGTCTI"

## ORIGIN

Query Match 86.2%; Score 3100.8; DB 4; Length 4493;  
Best Local Similarity 92.4%; Pred. No. 0;  
Matches 3296; Conservative 0; Mismatches 267; Indels 5; Gaps 3;  
15 GCGCGCAACTGAGCAAGGCTGAGGAGGAGGAGCGGCGGCGAGCGCTCGCGGCGAG 74  
249 GCGGCTGAGCGCGCGCGCGCGCGCTGCTGGTGTGGGCTCAGCGCGAGCGCA 308  
75 GAGCGGCGGCGCGCACTGTGTCAACGCGCGTGTGCTGCAACCTTGGCAACCGCGCGAGTGGCT 134  
309 GGAATGTGACCGGCGCTGTGTCAACGCGCGCGCTGTCAACCTTGGCAACCGCGCGAGTGTGT 368  
135 CTGCGCGCGCGCTGGGTGGCGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAAC 194  
369 CTGCGCGCGCGCTGGGTGGCGAGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAAC 428  
195 TGGATCTTCTGGATTTATAACAGATGACCTGGAAATTTATAAATAAATAAATAAATAAATAA 254  
429 TGGATCTTCTGGATTTATAACAGATGACCTGGAAATTTATAAATAAATAAATAAATAAATAA 488  
255 GTGGCTCATTTGAAGGACAGCAATAGAAATAGAGACTTGGTTTCAATCATTTTGGCTTAC 314  
489 ATGGCTCATTTGAAGGACAGCAATAGAAATAGAGACTTGGTTTCAATCATTTTGGCTTAC 548  
315 AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGGAATCAATTTATGCAACCGCTAGT 374  
549 AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGGAATCAATTTATGCAACCGCTAGT 608  
375 TGCTGCAATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGT 434  
609 TGCTGCTTTAGCGGCTCATTTGTTCTGAGAGAGAGCGCAATGAGACTGTCCAGAGT 668  
435 TGTGCAACATCAGTTTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 494  
669 TGTGCAACATCAGTTTATGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728  
495 TGAATTTAATTTACTTACAGTTTGTGATGTGTCOAATAAATGCTCAGCGCGAGGAGA 554  
729 TGAATTTAATTTACTTACAGTTTGTGATGTGTCOAATAAATGCTCAGCGCGAGGAGA 788  
555 GTGTAAGATCAGTAATAGCAGCATACTGTGTAATGTGTAATGTGTAATGTGTAATGTGTAATGT 614  
789 ATGTAAGATCAGTAATAGCAGCATACTGTGTAATGTGTAATGTGTAATGTGTAATGTGTAATGT 848  
615 TGAAGATCAGTAATCCTCATCTGTACAGCAACTGTGTTTTCCTCATCGAGGCACTGT 674  
849 TGAAGATCAGTAATCCTCATCTGTGTAAATAACTGTGTTTTCCTCATCGAGGCACTGT 908  
675 CAATTCAGTGATGTGAGGATGCTCTCTCTCTCAGACTGCGGCTCTGAGTGTTC 734



QY	2894	AGCCCCGTC	CAATTC	CAGCAT	TGTTCT	AGAGGAC	CAGCAT	CAACT	CGTCTTT	CAATTC	2953	
Db	3125	AGCCTCTT	CAATTC	CAGCAT	TGTTCT	AGAGGAC	CAGCAT	TAAC	TGGTCTTT	CAATTC	3184	
QY	2954	ACTGTCAG	CTTG	CAATG	CAAC	CGGCCAC	AGT	TAAT	CAATCA	ATCAGAGCAT	CTGT	3013
Db	3185	ATTGTCCAG	CTTG	CCAGT	GCAT	TGGCCAC	AGCAAT	GCAT	TAAAT	CAGAGTAT	CTGT	3244
QY	3014	AGTGTGAG	ACCT	GTGAC	CAAC	AGGCAAG	CACT	TGCGAG	AGCCT	GCATAT	CTGGCTT	3073
Db	3245	AGTGTGAG	AACCT	GTAC	CAC	CGGCGA	AGCACT	GTGT	GAGACCT	GCATAT	CTGGCTT	3304
QY	3074	ATCCCAACA	AT	TGGAGG	AAAT	GTCT	CAGCCAT	GTCA	AGT	TGGACCG	CGTCT	3133
Db	3305	ATCCAAACA	CGGGG	AAAT	GTCT	CAAC	CTGAGT	GCAT	TGGACAT	CGT	CGCTAT	3364
QY	3134	ACACCAAC	AC	CGGCGA	AGT	GTCTT	CTGCAC	CAAC	AGGCGCT	CAAGGGG	GCAGAGT	3193
Db	3365	ATACCAAC	AC	CGGCGA	AGT	GTCTT	CTGCAC	CAAC	AGGCGCT	CAAGGGG	GCAGAGT	3424
QY	3194	TATGTGAG	GTAG	AAAAAT	TCGAT	ACCA	AGGAAC	CCCT	CTCT	CAGAGGA	CAATGTT	3253
Db	3425	TATGTGAG	GTAG	AAAAAT	TCGAT	ACCA	AGGAAC	CCCT	CTCT	CAAGGA	CAATGTT	3484
QY	3254	TTCTTAT	TGACT	AT	CAGT	TTCA	CCCTTT	AGTCT	TAT	CCCCAG	GAAGTAT	3313
Db	3485	TTCTTAT	TGACT	AT	CAGT	TTCA	CCCTTT	AGTCT	TAT	CCCCAG	GAAGTAT	3544
QY	3314	CTATCAAT	TTTT	TGGCT	TACT	CTCT	GCACG	AAAC	CAAAAC	AGGGAAT	TGGACAT	3373
Db	3545	CCATCAAT	TTTT	TGGCG	AC	CCCT	GATG	AAAC	CAAAAC	AGGGAAT	TGGACAT	3604
QY	3374	CTCTCAAGA	ATTT	CAAC	CT	CAAC	AT	CACCT	GGGCT	GCAGT	TTCT	3433
Db	3605	CTCTCAAAA	ACTT	CAAC	CT	CAAC	AT	CACCT	GGGCT	GCAGT	TTCT	3664
QY	3434	CTGAGAGA	GAGAT	GCCT	TGTT	TTT	CAAAA	ACC	CAACAT	TAAGG	AGTAT	3493
Db	3665	CTGAGAGA	GAGAT	GCCT	TGTT	TTT	CAAAA	ACC	CAACAT	TAAGG	AGTAT	3724
QY	3494	CTAATGAG	AGATT	GTGAT	TTT	TGCGAAC	CAAC	CAACCAAT	ATAT	CATTTCT	TGTTTAT	3553
Db	3725	CTAATGAG	AGATT	GTGAT	TTT	TGCGAAC	CAAC	CAACCAAT	ATAT	CATTTCT	TGTTTAT	3784
QY	3554	ATTTCAC	CTCG	CCCAT	CAAAA	ATT	CAGGT	3581				
Db	3785	ATTTCAC	CTCG	CCCAT	CAAAA	ATT	CAGGT	3812				

RESULT 6

AF116897

LOCUS Mus musculus

DEFINITION Mus musculus mahogany protein mRNA, complete cds.

ACCESSION AF116897

VERSION AF116897.1

KEYWORDS GI:4454560

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 6083)

AUTHORS Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr., DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D., Fairchild-Huntress,V., Ge,P., Keilty,J., Ebeling,C., Baldini,L., Gilchrist,J., Burn,P., Carlson,G.A. and Moore,K.J.

TITLE The mahogany protein is a receptor involved in suppression of obesity

JOURNAL Nature 398 (6723), 148-152 (1999)

MEDLINE 93184159

PUBMED 10086355

REFERENCE 2 (bases 1 to 6083)

AUTHORS Nagle,D.L., McGrail,S.H., Vitale,J., Woolf,E.A., Dussault,B.J. Jr., DiRocco,L., Holmgren,L., Montagno,J., Bork,P., Huszar,D.,





1743 TGTGATAGTGAAGTGGACCATGCTGGTGTGTTGGAGGGAACACACAAATGACACTCCAT 1802  
1515 GAGCCATGGGCGCAAAATGCTTCTTCATGATTTTCATGGGCTATGACATTCCTGTGACCG 1574  
1803 GAGCCACGGTGCACAAATGCTTCTCTCGGACTTCATGGCTATGACATTCCTGTGACCG 1862  
1575 CTGGTCAAGTCTTCCAGACCTGA-TTCACCAATGATGTCACAGATTTGGCCATTCAG 1633  
1863 ATGGTCAAGTCTTCCAGACCTGAGCTCCATCATGATGTCACAGATTTGGCCATTCAG 1921  
1634 CAGTCTTACACACACACCAATGATGCTGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693  
1922 CAGTCTTACACACACCAATGATGCTGGTGGTTTCAATAGTCTCCTCCTCAGCG 1981  
1694 ACATCTGCTGATTCACCTGGACAGTGTGATGGCATCGGAGTGAAGCGCTGTTTAG 1753  
1982 ACCTCTTGGTCTTACCTCGGACAGTGGATGCACACCGCATGAGTGTCTGTGGG 2041  
1754 CAGCAGACCTGGTATTCCTGGTGTGTGGAACACAGGGTCTGCTCAGTGTATCTCGTGGG 1813  
2042 CAGCAGACCTGGTATTCCTGGTGTGTGGAACACAGTGTCTCGATGCTCCTCTGGG 2101  
1814 CGTGGCACTGATGAACAGAGAAAGTTAAATTCAGAAATGTTTTCCTCAAAAGATC 1873  
2102 AGTTGGCAATGAGAAACAGCAAGAAAGTTAAATTCAGAGTGTCTTCTAAAGAACCC 2161  
1874 TTGACCATGACATGTCGACGACACACAGATTTGTTACAGCTGCACAGGCAACCAATG 1933  
2162 TTGACCATGACATGTCGACGACACACAGATTTGTTACAGCTGCACAGGCAATCCATG 2221  
1934 ACTGCCACTGGTGCATGACCAATTTGTGTCGCCAGGAACACAGCTGCTCAGAGGCCAGA 1993  
2222 ACTGCCACTGGTGCATGACCAATTTGTGTCGCCAGGAACACAGCTGCTCAGAGGCCAGA 2281  
1994 TCTCCATTTTAGTATGAGATTTGCCCAAGATTAACCTATGACTCTGTACTGTACAGAG 2053  
2282 TCTCCATTTTAGTATGAGATTTGCCCAAGATTAACCTATGACTCTGTACTGTACAGAG 2341  
2054 AGACCACTGTCAGGAGCTGTGCCCTGGACACAGAACTGCGAGTGGGAGCCCGGAATCAGG 2113  
2342 AAACCACTGTCAGGAGCTGTGCCCTAGACACAGAACTGCGAGTGGGAGCCCGGAATCAGG 2401  
2114 AGTGCAATGCGCTGCGCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAATCAT 2173  
2402 AGTGCAATGCGCTGCGCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAATCAT 2461  
2174 GTTTGAAATTTACTCTGCAAGGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACC 2233  
2462 GTCTGAAATTTACTCTGCAAGGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACC 2521  
2234 ACAATGCCCTTTTGGCTTCTTTACAAACCAAGAGGATGAAATTTGTTCTTTAAGCAGC 2293  
2522 ACAATGCCCTTTTGGCTTCTTTACAAACCAAGAGGATGAAATTTGTTCTTTAAGCAGC 2581  
2294 TGGCAATATGACATCTCAGAGCATGTCAGAGCTCAACCTCACTTAACCCCATGGCTGGCC 2353  
2582 TTCGATTAATGCAATCATCTCAAGTATGTCACAGCTCACTCTGATCTCCATGGTGGTCTG 2641  
2354 TTCGGAAGTCAATGCTCTACTGCTGGGAAGATATGCCCATTTACAAATAGTT 2413  
2642 TTCGGAAGTCAATGCTCTACTGCTGGGAAGATATGCCCATTTACAAATAGTT 2701  
2414 TACTACAGTGGATGCGCTGAGCCAGTCAGTCATGCTGGATTTCTGTGGAATTTTATCAGAAC 2473  
2702 TGCTGAGTGGATGCGCATCTGAGCCAGTCAGTCATGCTGGATTTCTGTGGAATTTTATCAGAAC 2761  
2474 CCAGTACTCGGGGACTGGAAGCTGCAACCTGCAATCAACCCACTCAATGGTGGTCTGTG 2533  
2762 CTAGTACTCGGGGATTAAGGCTGCAACCTGCAATCAACCCCTCTCAATGGCAGCGTCTGTG 2821  
2534 AAAGGCTGCAACCAAGTCTGAGAGTGGCGGACACATGTCGCTTGGAGACAGCAT 2593  
2822 AAAGGCTGCAACCAAGTCTGAGAGTGGCGGACACATGTCGCTTGGAGACAGCAT 2881

RESULT 7  
AB062913  
LOCUS

AB062913 4284 bp mRNA linear ROD 16-JAN-2002

2594 GTGAGATTTGCAACAGCGGCAAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTG 2653  
2882 GTGCGGATGTGCACTAGCAGCAGCTCGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTG 2941  
2654 TGGATCTCAATGCTATGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2713  
2942 TGGACTCAATGCTATCGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3001  
2714 TGAGCACTGCCCCCTGAAATTTGTCAGCTACTGTACTGTAGTCTGATGCTGTTGAGC 2773  
3002 TGAGCAGTGGCCACCTGAAATTTGCTCTGCTACTGTACTGTAGTCTGATGCTGAGC 3061  
2774 AACCAAGCTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2833  
3062 AGCAGGCTGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3121  
2834 GTTCTTAAAGGACCAAGTGAAGTGTCTTGGCAAGCCCTTACAGGAAATTTCTATCCAC 2893  
3122 GCAGCTATAAGGACCTGTGAAGTGCCTCACAGGCTCTGCAAGGAAATGTGTATCCAC 3181  
2894 AGCCCTGCTCAATTCAGCATGTCTGAGGACAGCAGATACAACTGCTCTTCTTCTTCTTCT 2953  
3182 AGCCCTTCTGAATCTCAGCATGTCTGAGGACAGCAGATACAACTGCTCTTCTTCTTCTTCT 3241  
2954 ACTGTCCAGCTTGCATGCAACCGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGA 3013  
3242 ACTGTCCAGCTTGCATGCAACCGCCACAGTAAATGCAATCAATCAGAGTATCTGTGAGA 3301  
3014 AGTGTGAACCTGACCAAGGAAAGCACTGAGGACAGCAGATCTGCTGCTTCTGCTTCTTCT 3073  
3302 AGTGTGAAGCACTGACCAAGGAAAGCACTGAGGACAGCAGATCTGCTGCTTCTGCTTCTTCT 3361  
3074 ATCCCAACCAATGAGGAGAAATGTGAGCATGCAATGCAATGCAATGCAATGCAATGCAAT 3133  
3362 ACCGACTATGAGGAGAAATGTGAGCATGCAATGCAATGCAATGCAATGCAATGCAATGCA 3421  
3134 ACACCAACACCGGCAAGTGTCTTGACCAACCAAGGCGTCAAGGGGAGCAAGTCTTCTTCT 3193  
3422 ACACCAACACCGGCAAGTGTCTTGACCAACCAAGGCGTCAAGGGGAGCAAGTCTTCTTCT 3481  
3194 TATGTAGGTAGAAATTCGATACCAAGGAAACCTCTCAGAGGAACATGTTTATTTACTCT 3253  
3482 TATGTAGGTAGAAATTCGATACCAAGGAAACCTCTCAGAGGAACATGTTTATTTACTCT 3541  
3254 TTCTTATTTGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTTATTTACACAG 3313  
3542 TTCTCATTTGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTTATTTACACAG 3601  
3314 CTATCAATTTTGTGGCTACTCTGACGAAACCAACAGGATTTGGACATGTTTATCAATG 3373  
3602 CCATCAATTTTGTGGCTACTCTGATGACAAACAGGATTTGGACATGTTTATCAATG 3661  
3374 CCTCAAGATTTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433  
3662 CCTCAAGATTTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3721  
3434 CTGAGAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3493  
3722 CTGAGAGAGATGCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCT 3781  
3494 CTAATGAGAGTGTGATTTTGGCAACCAACCAATATCACTTCTTCTTCTTCTTCTTCTTCT 3553  
3782 CTAATGAGAGTGTGATTTTGGCAACCAACCAATATCACTTCTTCTTCTTCTTCTTCTTCT 3841  
3554 ATTTCACTGGCCCATCAAAATTCAGGT 3581  
3842 ATTTCACTGGCCCATCAAAATTCAGAT 3869



















Db 486 TGGATCTTCTGGATTGTATACAGATGGACCTGGGAATTATAAATATATAAACAAGAAATGCAC 545  
QY 255 GTGGCTCAITTAAGAGCAGCAAAATAGATAATATGAGACTTCGTTTCAATCAATTTGCTAC 314  
Db 546 ATGGCTCAITTAAGAGCAGCAAAATAGATAATATGAGACTTCGATTCAACCAITTTGCTAC 605  
QY 315 AGAGTGTAGTTGGACCAATTAATATATGTTATGATGGGACTCAATTTATGCAACCGTAGT 374  
Db 606 TGAATGTAGCTGGACCAATTAATATGTTATGATGGGACTCAATTTATGCAACCTCTGAT 665  
QY 375 TGTGCAATTTAGTGGCTCAATTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434  
Db 666 TGTGCTCTTTAGTGGCTCAATTTCTGAGAGAGATGGCAATGAGACGCGCCCTGAACT 725  
QY 435 TGTGTCACATCAGTTATGCTTGTGCAATTTTATGATGGGCTTATGATTTGAT 494  
Db 726 CACTGTCACTTCAAGTTATGCATCTGCAATTTTTCAGTGATGCTTATTAATCTGAC 785  
QY 495 TGGATTTAATTAATTAATTAATTTGATATGTTCCAAATAACTGCTCAGGCGGAGAGA 554  
Db 786 TGGATTTAATTAATTAATTTGACATGTGCCGNAATAATTTGCTCAGGCGGAGAGA 845  
QY 555 GTGTAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAATCTGAAAGG 614  
Db 846 ATGTAAGAGCAGTAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAATCTGAAAGG 905  
QY 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGTTTTCCTCATCGAGGCATCTG 674  
Db 906 GGAATCTGTGACATTCCTCACTGTACAGACAACTGTGCTTTCCTCACCAGGCAATCTG 965  
QY 675 CAATTCAGATGATGTGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCTGGATGTTTC 734  
Db 966 TAAACGAAGTGAACACAGAGGGTCTCCTGCTTCCCTCACTGGCAGGGTCTGGATGTTTC 1025  
QY 735 AGTTCTCTACAGCTAAACAGTCAATTTGGACTCGAGAGGAATATTTAACTTAAAGCT 794  
Db 1026 AATTCCTGTGCCAGCTTAACAGTCTTTTGGACTCGAGAGGAATATTTCTGACTTAAAGCT 1085  
QY 795 CCCAGAGCATCTCATAAAGCTGTGTCATATGTAAGAAACATTTATGTGGTTGTTGGAGATA 854  
Db 1086 TCCAGAGCCTCTCATAAAGCTGAGGTCAATGGAAATAATAATGTGGTTGTTGGCGGTA 1145  
QY 855 TATGTTCAACCACTCAGATTATAACATGTTCTAGCGTATGACCTGCTTCTAGGAGTG 914  
Db 1146 TATGTTCAACCACTCAGATTATACAGCATGTTTATAGCTATGACCTGGCTTCTAGGGAATG 1205  
QY 915 GCTTCCACTAAACCGTCTGTGAAACAAATGTGGTTGTTAGATATGGTCAATCTTTGGCAAT 974  
Db 1206 GCTTTCACTTAACCAATCTGTGAAACAGTGTGGTTGTTAGATATGGTCAATCTTTGGCAAT 1265  
QY 975 ATACAAGGATAAAATTTAATGTATGGAGAAATAATGATTCAACTGGGAATGTGACCAA 1034  
Db 1266 ACATAAGGATAAAATCTACATGTATGGAGAAATAATGATTCAACAGGGAATGTGACCAA 1325  
QY 1035 TGAGTTGAGAGTTTTCACATTCATAATGAGTATGAGTGTGTTGACCTCAAGCCCTTAAGGCAA 1094  
Db 1326 CGAGCTGAGAGTATTTCAATTCATAATGAATCAATGGGATTTGTTAAACCCCAAGCTAA 1385  
QY 1095 GGAGCAGTATGCAAGTGTGGGCACTTCGCACACATTTGTACACTGAAGAAATGGCCGAGT 1154  
Db 1386 GGATCAGTATGCAAGTGTGGGCACTTCGCACACATTTGTACACTGTCTATCTGGCCGTGT 1445  
QY 1155 GGTTCATGCTGGTCAATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGA 1214  
Db 1446 GGTCAATGTTGGTCAATCTTTGGTCACTGCCCTCTACGGATATATAAGTGTGTTGCAGGA 1505  
QY 1215 ATATGATTTGGATAAGAAACACATGGAGTATATTAACACCCAGGGTGCCTTGTGCAAGG 1274  
Db 1506 ATATGACTTTGGAAAGAAACACATGGAGTATATTAACAACTCAAGGTGCTCTAGTCAAGG 1565  
QY 1275 GGGTTACGGCATAGCAGTGTGTTACCAACATAGGACCGGGCCCTATAGCTTCATGGTGG 1334  
Db 1566 GGGTTATGCGCCACAGCAGTGTGTTATGATCACAGGACCAAGGCTCTGTATGTTTCATGGTGG 1625

QY 1335 CTACAAGGCTTTCAAGTGCACAATAAGTACCGGCTTGACAGATGATCTCTACCAATATGATGT 1394  
Db 1626 CTACAAGGCTTTCAAGTGCACAATAAGTACCGGCTTGACAGATGATGATGATGATGATGAT 1685  
QY 1395 GGATACCCAGAGATGTGGACCAATCTTTAAGGACAGCGGATTTTCCGTTACTTGGCACACAG 1454  
Db 1686 GGACACTCAGATGTGGACCAATCTTTAAGGACAGCGGATTTTCCGTTACTTGGCATACAGC 1745  
QY 1455 TGTGATGATGATGGAAACCAATGCTGTTGTTGGGGGAAACACACAAATGACACACTAT 1514  
Db 1746 TGTGATGATGATGGAAACCAATGCTGTTGTTGGGGGAAACACACAAATGACACTTCCAT 1805  
QY 1515 GAGCATTGGCCCAAAATGCTTCTCTCAGATTTTATGSCCTTATGACATTTGCTGACCG 1574  
Db 1806 GAGCCACGGTGCACAAATGCTTCTCTCAGACTTATGSCCTTATGACATTTGCTGACCG 1865  
QY 1575 CTGCTCAGTGTCTCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCATTTCCAG 1633  
Db 1866 ATGCTCAGTCTTCCAGACCTGAGCTCCATCAAGATGTCAACCGATTTGGCCACT-CAG 1924  
QY 1634 CAGCTTTACACACAGACCAATGATGTTGCTGGTGTTCATAGTCTTCAATAGTCTTCTCTCAGCG 1693  
Db 1925 CAGCTTTGACACACAGACCAATGATGTTGTTGCGGCGCTTCAACAGCGCTTCTCTCAGTG 1984  
QY 1694 ACATCTGCTATTCACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTGTAG 1753  
Db 1985 ACGCTTAGTCTTCACTCGAGCAGTGTGATGACACCGCAGCGAAGCTGCTTGTGTGG 2044  
QY 1754 CAGCAGGACCTTGGTATTTGCGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGG 1813  
Db 2045 CAGCAGGACCTTGGCATCCGGTGTCTGTGGGACACACAGTCTGCTCGATGTACCTCTCTGGG 2104  
QY 1814 CGCTGGCAACTGATGAACAAGAAGAAAGTTAAATCAGATGTTTCTTAAAGAAACCC 2164  
Db 2105 AGTTGGCAACTGAAGAACACAGCAAGAAAGTTAAATCAGATGTTTCTTAAAGAAACCC 2164  
QY 1874 TTGACCATGACAGATGTGACCAACACAGATTTGTTTACAGCTGCACAGCCCAACCAATG 1933  
Db 2165 TTGACCATGACAGATGTGACCAACACAGATTTGTTTACAGCTGCACAGCCCAATACCAATG 2224  
QY 1934 ACTGCCACTGGTGAATGACCAATGTTGTGCTCCCAAGAACCAAGCTGCTCAGAGAGCCAGA 1993  
Db 2225 ACTGCCACTGGTGAATGATCACTGTGCTCCCTGTGAAACCAAGCTGCAAGAGCCAGA 2284  
QY 1994 TCTCCATTTTATGATGATGAGAAATGCCCCAAGGATAACCTTATGCTACTGTAAACAAGA 2053  
Db 2285 TCTCCATTTGCAAGTATGACAAATTTGCCCAAGGATTAACCCATGATCTACTGTCAATAGA 2344  
QY 2054 AGACAGCTGCAGAGCTGTGCTCCCTGGACCAAGAACTGCCAGTGGGAGCCCGGGAATCAGG 2113  
Db 2345 AAACCAAGCTGCAGAGCTGTGCTCCCTAGATCAGAAATTCAGATGGGAACTCGAAATCAAG 2404  
QY 2114 AGTCATTTGCCCTGCCGAAATATCTGTGCAATGGCTGGCAATTTGCTGGTGGAACTCAT 2173  
Db 2405 AGTGCATCGCCCTGCCGAAACATCTGTGCAATGGCTGGCAATTTGCTGGTGGAACTCAT 2464  
QY 2174 GTTTGAAATTAATCTACTGCCAAGGAAATTAATGCAATGCTTAAATTTGCTGTAGGAACC 2233  
Db 2465 GTCTGAAATTAATCTACTGCCAAGGAAATTAATGCAATGCTTAAATTTGCTGTAGGAACC 2524  
QY 2234 ACAATGCCCCCTTTGGCTTCTTACAAACCCAGAAAGGTAGAAATTTGCTCTTAAAGCAGC 2293  
Db 2525 ACAATGCCCCCTTTGGCTTCTGCTTACGCTCCAGAAAGGTGGAGTTTGTCTTAAAGCAGC 2584  
QY 2294 TGGCAATTAATGCACTCATCTCAGAGCATGTCCAAGCTCACTTAAACCCCTGGGTCCGCC 2353  
Db 2585 TTCCGATTAATGCAATCATCTCAGAGCAGCTCCAAGCTCACTCTGACTCCATGGGTGGTTC 2644  
QY 2354 TTCCGAGATCAATGCTCCTAATGCTGCTGGGAGGATATGTCGCCCATTTTACAAATAGTT 2413  
Db 2645 TTCCGAGATCAATGCTCCTAATGCTGCTGGGAGGATATGCTCTCCATTTACAAATAGTT 2704

2414 TACTACAGTGGATGCGGTCTGAGCCGAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2473  
Db TGCTGAGTGGATGCCATCTGAGCCGAGTGATGCTGGATTCTGTGGAATTTTATCAGAAC 2764  
2474 CCAGTACTCGGGGACTGAGGGCTGCAACCTGCAATCAACCACTCAATGATGATGCTGTG 2533  
Db CTAGTACTCGGGGACTGAGGGCTGCAACCTGCAATCAACCACTCAATGATGATGCTGTG 2824  
2534 AAAGGCTGCAAAACCAACAGTGTGAGGAGTGGCGGACACCATGCTGCTTGAGGACAGCAT 2593  
Db AAAGGCTGCAAAACCAACAGTGTGAGGAGTGGCGGACACCATGCTGCTTGAGGAGTGG 2884  
2594 GTGGAGATTCACAGGAGGAGTCTGAGTGCATGTGGTGCACCAATCAAGAGCAGTGTG 2653  
Db GTGGAGATTCACAGGAGGAGTCTGAGTGCATGTGGTGCACCAATCAAGAGCAGTGTG 2944  
2654 TGGAATCCAAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGA 2713  
Db TAGACTCAATGCTATGCTAGTCTCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGA 3004  
2714 TGAGCACTGCGCCCTGAAAATTTGTCAGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTG 2773  
Db TGAGCACTGCGCCCTGAAAATTTGTCAGGCTACTGCTGCTGCTGCTGCTGCTGCTGCTG 3064  
2774 AACGAGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2833  
Db AACGAGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3124  
2834 GTTCCTATAAGGACCAAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2893  
Db GTTCCTATAAGGACCAAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3184  
2894 AGCCCTGCTCAATTCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2953  
Db AGCCCTGCTCAATTCAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3244  
2954 ACTGTCCAGCTTGCCAAATGCAAGGCGCACAGTAAATGCAATGCAATGCAATGCAATGCA 3013  
Db ACTGTCCAGCTTGCCAAATGCAAGGCGCACAGTAAATGCAATGCAATGCAATGCAATGCA 3304  
3014 AGTGTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3073  
Db AGTGTGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3364  
3074 ATCCCAACCAATGGAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3133  
Db ATCCCAACCAATGGAAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3424  
3134 ACACCAACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3193  
Db ACACCAACCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3484  
3194 TATGTGAGTGTGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3253  
Db TATGTGAGTGTGAGAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3544  
3254 TTCTTTATGATATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3313  
Db TTCTTTATGATATCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3604  
3314 CTATCAATTTTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3373  
Db CTATCAATTTTGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3664  
3374 CCTCAAGAAATTCACCTCAACATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3433  
Db CCTCAAGAAATTCACCTCAACATCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3724  
3434 CTGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3493  
Db CTGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3784  
3494 CTAATGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3553

3785 CTAATGAGAAATTTGATTTTTCGCAACCATCAAACTACTTCTTCTTGTATGTCAGTA 3844  
QY 3554 ATTTCACTGGCCCATCAAAATTCAGGT 3581  
Db 3845 ATTTCACTGGCCCATCAAAATTCAGGT 3872  
RESULT 11  
ARI164807 8827 bp DNA linear PAT 17-OCT-2001  
LOCUS Sequence 1 from patent US 6274339.  
DEFINITION ARI164807  
ACCESSION ARI164807  
VERSION ARI164807.1 GI:16238010  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 8827)  
AUTHORS Moore, K. and Nagle, D. Lynn.  
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity  
JOURNAL Patent: US 6274339-A 1 14-AUG-2001;  
FEATURES Location/Qualifiers  
1..8827  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Query Match 79.1%; Score 2845.2; DB 6; Length 8827;  
Best Local Similarity 90.3%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 328; Indels 2; Gaps 2;  
QY 189 ACTAACTGAGTCTTCTGGGTTTGTGACAGATGGACCTGGAAATATATAATACAAAGCAA 248  
Db 677 ACTAACTGAGTCTTCTGGGTTTGTGACAGATGGACCTGGAAATATATAATACAAAGCAA 736  
QY 249 GTGACAGTGGCTCAATGAGGACGACCAATAGAAATAGAGCTTCGTTCAATCAATTT 308  
Db 737 GTGACAGTGGCTCAATGAGGACGACCAATAGAAATAGAGCTTCGTTCAATCAATTT 796  
QY 309 TGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACCTCAATTTATGCACC 368  
Db 797 TGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACCTCAATTTATGCACC 856  
QY 369 GCTAGTTCCTGCAATTAGTGGCTCTATGTTCTGAGAGATGGCAATGAGACTGCC 428  
Db 857 TCTGATTCCTGCTTCTAGTGGCTCTATGTTCTGAAAGAGATGGCAATGAGACTGCC 916  
QY 429 TGAGTTCCTGCAATTAGTGGCTCTATGTTCTGAGAGATGGCAATGAGACTGCC 488  
Db 917 TGAGTTCCTGCAATTAGTGGCTCTATGTTCTGAGAGATGGCAATGAGACTGCC 976  
QY 489 TTTGACTGGATTTAATATTACTTACAGTTTGTATGTTCCCAATTAATCTGCTCAGGCCG 548  
Db 977 TCTGACTGGATTTAATATTACTTACAGTTTGTATGTTCCCAATTAATCTGCTCAGGCCG 1036  
QY 549 AGGAGAGTGAAGATCAGTAAATAGCAGGATGTTGTAATGTTGTTCTGAAATCTG 608  
Db 1037 AGGAGAGTGAAGATCAGTAAATAGCAGGATGTTGTAATGTTGTTCTGAAATCTG 1096  
QY 609 GAAAGGTGAAGCATGTGACATTCCTCTGCTACAGACAACTGTGGTTCCTCATTCAGG 668  
Db 1097 GAAAGGTGAAGCATGTGACATTCCTCTGCTACAGACAACTGTGGTTCCTCATTCAGG 1156  
QY 669 CATCTGCAATTCAGTGAATGTCAGAGGATGCTCTGCTCTCAGACTGGCAGGCTCTCG 728  
Db 1157 CATCTGCAATTCAGTGAATGTCAGAGGATGCTCTGCTCTCAGACTGGCAGGCTCTCG 1216  
QY 729 ATGTTTCAGTTCCTGTACAGCTAACAGTCAATTTTGGACTCGAGGAAATTTCTAATTT 788  
Db 1217 ATGTTTCAGTTCCTGTACAGCTAACAGTCAATTTTGGACTCGAGGAAATTTCTAATTT 1276

789 AAAGCTCCCGAGAGCATCTCATAAAGCTGTGTCATGGAACATTATATGTGGGCTTTGG 848  
1277 AAAGCTTCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAATAATATGTGGGCTTTGG 1336  
849 AGGATATATGTTCAACCACTCAGATTATAACATGGTCTTAGCGTATGACCTTGCTTCTAG 908  
1337 CGGATATATGTTCAACCACTCAGATTACAGCATGGTTTTAGCGTATGACCTGACTTCTAG 1396  
909 GGAGTGGCTTCCACTAAACCGTCTGTGGAACATGGTTGTTAGATATGGTCAATTCITT 968  
1397 GGAATGGCTTCCACTAAACCACTCTGTGGAACATGGTTGTTAGATATGGTCAATTCITT 1456  
969 GGCATTATACAAGGATAAATAATTTACATGATGAGGAAATAATGATTTCAACTGGGAATGT 1028  
1457 GGCATTATACAAGGATAAATAATTTACATGATGAGGAAATAATGATTTCAACGGGAACGT 1516  
1029 GACCAATGAGTGTGAGAGTTTTTACATTCATATGATGATGATGATGATGATGATGATGAT 1088  
1517 GACCAATGAGTGTGAGAGTTTTTACATTCATATGATGATGATGATGATGATGATGATGATGAT 1576  
1089 GGCAAAGGAGCATGATGACGTGGTGTGGGCACTCTGCACACATTTTACACTGAAGAATGG 1148  
1577 AGCTAAGGATCAGTATGCACTGGTGTGGCACTCAGCACATTTGTTACATGCGCATCTGG 1636  
1149 CCGAGTGGTCAATGCTGGTCACTTTTGGTCACTGCTGCTCTCTATGGAATATAAGCAATGT 1208  
1637 CCGTGTGGTCAATGCTGGTCACTTTTGGTCACTGCTGCTCTCTATGGAATATAAGCAATGT 1696  
1209 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATACACACCCAGGGTGCCTTGT 1268  
1697 GCAGGAATATGATTTGGATAAGAACACATGGAGTATATACACACCCAGGGTGCCTTGT 1756  
1269 GCAAGGGGTTACGGCCATAGCAGTGTGTTACGACCATAGACACGAGGCGCCCTATACGTTCA 1328  
1757 GCAAGGGGTTATGGCCACAGTATGTTTATGATGACAGCAACGCTCTGTACGTTCA 1816  
1329 TGTGGCTACAAAGCTTTCAGTGGCAATAGTACCGCTTGCAGATGATCTCTACCGATA 1388  
1817 TGTGGCTACAAAGCTTTCAGTGGCAATAGTACCGCTTGCAGATGATCTCTACAGATA 1876  
1389 TGATGTGGATACCAAGTGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTTGCA 1448  
1877 CGATGTGGATACCAAGTGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTTGCA 1936  
1449 CACAGCTGTATAGTGTGAGTGAACCAATGCTGTGGTTTTTGGGGGAAACACACAAATGACAC 1508  
1937 TACAGCTGTATAGTGTGAGTGAACCAATGCTGTGGTTTTTGGGGGAAACACACAAATGACAC 1996  
1509 ATCTATGAGCATGGCGCCAAATGCTTCTTTCAGATTTTCATGGCCCTATGACATGCGCTG 1568  
1997 TTCCATGAGCCACGGTGCMAATGCTTCTCTCGGACTTCATGGCTTATGACATGCTTGT 2056  
1569 TGACCGCTGTGCTAGTCTTCCAGACCTGA - TTCCACCATGATGTCAACAGATTTTGGCCA 1627  
2057 TGACCGATGTGCTAGTCTTCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116  
1628 TTCCAGAGCTTTACACACAGCACCATGATGTGTTCCGTTGTTTCAATAGTCTCTCC 1687  
2117 TT - CAGCAGTCTTGTACACAGCACCATGATGTGTTCCGCGGCTTCAACAGCCTCTCTCC 2175  
1688 TCAGCAGCATCTCTGGTATTCACTCCGAAACAGTGTGATGCGCATCGGAGTGAAGCGCTT 1747  
2176 TCAGTGAAGCTTCTTGGTCTTACCTCCGAGCAGTGCATGCAACCGCAGTGAAGCTGCTT 2235  
1748 GTTTAGCAGCAGACCTGTATTCGGTGTGTTGGAAACACAGGTCGTCTCAGTGTATCT 1807  
2236 GTTGGCAGCAGACCTGTATTCGGTGTGTTGGAAACACAGGTCGTCTCAGTGTATCT 2295  
1808 CGTGGCGCTGGCAACTGTATGAACAGAGAAAGTTAAATCAGATGTTTTTCCAAA 1867  
2296 CTTGGAGTGTGGCAACTGTATGAACAGAGAAAGTTAAATCAGATGTTTTTCTAAA 2355  
1868 GAATCTTTGACCATGACAGATGTGACCGACACACAGATTTGTTACAGCTGCACAGCCACA 1927

2356 GAAACCTTTGACCAATGACAGATGTGACGAGCACACAGATTTTACAGCTGCACAGCAATA 2415  
1928 CCAATGACTGCCACTGTGTCAATGACCAATTTGTGTCCCGAGGAACACACAGCTGCTCAGAA 1987  
2416 CCAATGACTGCCACTGTGTCAATGACCAATTTGTGTCCCGAGGAACACACAGCTGCACAGA 2475  
1988 GCCAGATCTCCATTTTGTAGTATGAGATTTGCCCGAGGATAACCTATGTACTACTGTA 2047  
2476 GCCAGATCTCCATTTGCAAGTATGAGATTTGCCCGAGGATAACCTATGTACTACTGTA 2535  
2048 ACAAGAACACAGCTGCAGGAGCTGTGCTTGGACCCAGAACTGCCAGTGGGAGCCCGGA 2107  
2536 ATAAGAAACACAGCTGCAGGAGCTGTGCTTGGACCCAGAACTGCCAGTGGGAGCCCGGA 2595  
2108 ATCAGGAGTGCATTTGCTGCTGCCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGA 2167  
2596 ATCAGGAGTGCATTTGCTGCTGCCCGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTGA 2655  
2168 ACTCATGTTTGAATAATTTACTGTGCAAGGAGAAATATGACAAATGCTAAATTTTCTGTA 2227  
2656 ACTCGTGTCTGAATAATTTACTGTGCAAGGAGAAATATGACAAATGCTAAATTTTCTGTA 2715  
2228 GGAACCAATGCTGCTTTTGGCTTCTTCAACCCAGAAAGAGTAAATTTTCTGCTTA 2287  
2716 GGAACCAATGCTGCTTTTGGCTTCTTCAACCCAGAAAGAGTAAATTTTCTGCTTA 2775  
2288 AGCAGCTGCGAATTAATGAGTCACTCAGAGCATGTCACAGCTCAACCTTAACCCCATGG 2347  
2776 AGCAGCTGCGAATTAATGAGTCACTCAGAGCATGTCACAGCTCAACCTTAACCCCATGG 2835  
2348 TCGGCTTTGGAAGATCAATGTGCTTACTGTGCTGCGAAGATATGCTCCCATTTTACAA 2407  
2836 TTGGTCTTTGGAAGATCAATGTGCTTACTGTGCTGCGAAGATATGCTCCCATTTACAA 2895  
2408 ATAGTTTACTACAGTGTGCTGCTGAGCCAGTGTGCTGGAATTTCTGCGAATTTTAT 2467  
2896 ATAGTTTCTCAGTGTGCTGCTGAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2955  
2468 CAGAACCCAGTACTCGGGGACTGAAAGCTGCAACCTGCTCAACCCACTCAATGCTAGTG 2527  
2956 CAGAGCTGTACTCGGGGATTAAGGCTGCAACCTGCTCAACCCACTCAATGCTAGTG 3015  
2528 TCTGTGAAGGCTTGCAGAAACACAGTGTGCTGAGGCTGCGGACACCATGTCCTTGAGGA 2587  
3016 TCTGTGAAGGCTTGCAGAAACACAGTGTGCTGAGGCTGCGGACACCATGTCCTTGAGGA 3075  
2588 CAGCATGTGGAGATTGCACACGCGCTGCTGAGTGCATGCTGCTGAGCAACATGAAGC 2647  
3076 CAGCGTGTGGAGTGCATGAGCAGCTGCGAGTGCATGCTGCTGAGCAACATGAAGC 3135  
2648 AGTGTGGAATCCATGCTGCTTATGCTGCTTCTTCCCTTTTGGCAGTGTATGGAATGGT 2707  
3136 AGTGTGGAATCCATGCTGCTTATGCTGCTTCTTCCCTTTTGGCAGTGTATGGAATGGT 3195  
2708 ATAGATGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2767  
3196 ATAGATGAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3255  
2768 TGGAGCAACAGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2827  
3256 TGGAGCAACAGGCTGTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3315  
2828 TAGAGGTTCTTAAAGGACAGTGAAGTGTGCTTCCGAGCCCTTACAGGAAATTTCT 2887  
3316 TTGAGGAGCAGTAAAGGACCTGTGAAGTGTGCTTCCGAGCCCTTACAGGAAATTTCT 3375  
2888 ATCCACAGCCCTCTCAATTTCCAGCATGCTGTAGAGGACAGCAGATACAACTGCTT 2947  
3376 ATCCACAGCCCTCTCAATTTCCAGCATGCTGTAGAGGACAGCAGATACAACTGCTT 3435  
2948 TCATTCTGCTCAGCTTGGCAATGCAACGCGCCACAGTAAATGCAATCAATCAGAGCATCT 3007

Db 3436 TCATTCACTGTCAGCTGCCAGTGCACCGACACAGCAAAATGCATCAACCGAGATATCT 3495

Qy 3008 GTGAGAAAGTGTGAGAACTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3067

Db 3496 GTGAGAAAGTGTGAGAACTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3555

Qy 3068 ACGGTATCCCAACAATGAGAGGAATGTGAGGCAATGTCAGGCAATGCAAGTCAATGGCAGCGTCTC 3127

Db 3556 ATGTTGACCCGACTAATGAGGCAATGTGAGGCAATGTCAGGCAATGCAAGTCAATGGCAGCGATCAC 3615

Qy 3128 TGTGCAACACCAACAGGCAAGTGTCTGACACCAAGGCGCTCAAGGCGGACGAGT 3187

Db 3616 TGTGCAACACCAACAGGCAAGTGTCTGACACCAAGGCGCTCAAGGCGGACGAGT 3675

Qy 3188 GCCAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAATGTTAT 3247

Db 3676 GCCAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAATGTTAT 3735

Qy 3248 ATACTCTTCTATGACTATGACTTACCTTTAGTCTATCCAGGAGGATGATCGCTATT 3307

Db 3736 ATACCTCTCTATGACTATGACTTACCTTTAGTCTATCCAGGAGGATGATCGCTATT 3795

Qy 3308 ACACAGCTATCAATTTTGTGGTACTCTCTGACGAAACAAACAGGATTTGGACATGTCA 3367

Db 3796 ACACAGCTATCAATTTTGTGGTACTCTCTGACGAAACAAACAGGATTTGGACATGTCA 3855

Qy 3368 TCAATGCTCAGAAATTTCAACTCAATCACTGGGTCGAGTTCCTCAGCTGGAA 3427

Db 3856 TCAATGCTCAGAAATTTCAACTCAATCACTGGGTCGAGTTCCTCAGCTGGAA 3915

Qy 3428 CCCAGGCTGAGAGAGATGCTTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATA 3487

Db 3916 CCCAGGCTGAGAGAGATGCTTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATA 3975

Qy 3488 GTTCTCTAATGAGAACTTTGATTTGCGAACCAACCAATATCACTTTCTTTGTTATG 3547

Db 3976 GTTCTCTAATGAGAACTTTGATTTGCGAACCAACCAATATCACTTTCTTTGTTATG 4035

Qy 3548 TCAGTAATTTTCACTGGCCCATCAAAATTCAGT 3581

Db 4036 TCAGTAATTTTCACTGGCCCATCAAAATTCAGT 4069

RESULT 12

AB093245

LOCUS

DEFINITION Mus musculus mRNA for mKIAA0548 protein.

ACCESSION AB093245

VERSION AB093245.1 GI:26006172

KEYWORDS FLI CDNA.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1

AUTHORS Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Hara, Y., Nagase, T., Ohara, O. and Koga, H.

TITLE Prediction of the coding sequences of mouse homologues of KIAA gene: 1. The complete nucleotide sequences of 100 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 5683)

AUTHORS Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.

TITLE Direct Submission

JOURNAL Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: mouse@kazusa.or.jp, tel:81-438-52-3919, fax:81-438-52-3918)

COMMENT The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing; Kazusa DNA Research Institute; cDNA library construction; clone selection and 5'- & 3'-end one pass sequencing.

FEATURES

source

Location/Qualifiers

1. 5683

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="mbg03041"

/tissue\_type="brain"

/dev\_stage="adult"

/note="vector:modified pBC SK-"

1. 5683

/gene="mKIAA0548"

<1. 3181

/gene="mKIAA0548"

/note="CDS is predicted by in silico analysis. Start codon is not identified."

/codon\_start=2

/evidence=not\_experimental

/product="mKIAA0548 protein"

/protein\_id="BAC41429.1"

/db\_xref="GI:26006173"

/translation="AYDLTSREMLPLNHSVNSVVRVGHSLALHDKIKVYGGKIDLVG  
GNVTVRPHIHNESWVLT PKAQOYAVVGHSAHIVTLASGRVLMVIFGHCPYLVG  
YISVQGEIDLEKNTWSILHTQGAALVQGGYSHSVYDDRTKALVHGYGAFSAKXRL  
ADDLYRYDVTQMTILKSRFPYRHLTVIVSGTLMVYFGNFNSLLSDLVLFLE  
DFMAYDIACDRWSVLPPELHVDNRFHSAVLYNTMYVFGNFNSLLSDLVLFLE  
QCAHRSFAACAAAGPGRICLWDTSRSTSWELATEBAEKLSKCFSKRTLDHRC  
DOHTDCYSTANTNDCHNDHCVPNHSCTEGQISIAKYSPKDPMPYCNKHTSC  
RSCALDONCOWEPNOECIALPENICGNHGLVNSCLKITAKENYDNALSKRNH  
AFASLTSQKYEPVLQBLRMOSSOSMSKLTLPVGLKIKINYSVWEDMSPFTNS  
LLQMPSEPDAGFCGLSEPSRGLKAATCINPLNGSVCEPANSKOCPTQALR  
TACBECTSSSECCMSNKCVDNAYVASFPFGQCMWITMSSCPPENCSCYTCGS  
HCLPQCGCTDPSNTGKGCIEGYSYGFVPMPSQASAGNVYPQLNSMCLDSR  
YNGFIFHCPCQCNCNKINQSVCECDLTGKHCETCISGFYDPTNGKCPCK  
CNGHSLCNTNTGKCTTKGVKDEQCLCEVENRYQGNPLGTCYTTLLIDYFTFS  
LSQDDRYTALNFVATPDQRDLDMFINAKNFNLNIWTATSPAGTGTGEVFPV  
SKTTIKYKDSFSENEKDFENHNTIFVVSNTFWPKIKIIFASHSNFMDLVQFV  
FTTFCFLSLLVAQAAWKKIKQSCWASRRRQLLEMOQASRPASVNVVALETDESP  
DLIGSITKVPKPIALEPCFNGKAAVLVSFVRLPRGLGGIPFPQSGSLAVASALVDS  
QQMPIVYKESGAVNRNRKQPPAPGTCI"

ORIGIN

Query Match 62.7%; Score 2255; DB 10; Length 5683;

Best Local Similarity 90.3%; Pred. No. 0;

Matches 2433; Conservative 0; Mismatches 260; Indels 2; Gaps 2;

Qy 888 AGCGTATGACCTTCTTCTAGGAGTGGCTTCCACTAAACGGTCTCTGTGAACAATGTGGT 947

Db 1 AGCGTATGACCTTCTTCTAGGAGTGGCTTCCACTAAACGGTCTCTGTGAACAATGTGGT 60

Qy 948 TGTAGATATGCTCATCTTTGGCATTATCAAGGATAAAATTTACATGTATGAGGAAA 1007

Db 61 TGTAGATATGCTCATCTTTGGCATTATCAAGGATAAAATTTACATGTATGAGGAAA 120

Qy 1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTGTTCACATTCATTAATGATG 1067

Db 121 AATTGATTCAACAGGACGTCACCAATGAGTGTGAGAGTGTTCACATTCATTAATGATG 180

Qy 1068 ATGGTGTGTTGACCCCTAAGCAAGGACAGTATGCGAGTGTGGGCACTCTGCACA 1127

Db 181 ATGGTGTGTTGACCCCTAAGCAAGGACAGTATGCGAGTGTGGGCACTCTGCACA 240

Qy 1128 CATTGTTACACTGAAGAATGGCCGAGTGGTCAATGCTGTCATCTTGGTCACTCCCTCT 1187

Db 241 CATTGTTACACTGGCATCTGGCCGTTGGTGTCAATGCTGTCATCTTGGTCACTCCCTCT 300

Qy 1188 CTTATGGATATATAGCAATGTGACCAATATGATTGGATTAAGACACATGCGATATATT 1247

Db 301 CTTATGGATATATAGCGTGTGTGACCAATATGATTGGATTAAGACACATGCGATATATT 360

Qy 1248 ACACACCCAGGTCCTTGTGCAAGGGGGTGTAGCGGCATAGCAGTGTGTTACGACCATAG 1307

Db 361 ACATCTACTAGGTCCTTGTGCAAGGGGGTGTAGCGGCATAGCAGTGTGTTATGATGACAG 420



Db	2580	CCATCAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAA	2639
Qy	3527	ATATCACTTTCTTTGTTTATGTACGTAAATTTACCTGGCCCATCAAAATTCAGGT	3581
Db	2640	ACATCACTTTCTTTGTTTATGTACGTAAATTTCACTTGGCCCATCAAAATTCAGAT	2694
RESULT 13			
LOCUS	AR164818	2625 bp	DNA linear PAT 17-OCT-2001
DEFINITION	Sequence 18 from patent US 6274339.		
ACCESSION	AR164818		
VERSION	AR164818.1	GI:16238029	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2625)		
AUTHORS	Moore, K. and Nagle, D. Lynn.		
TITLE	Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity		
JOURNAL	Patent: US 6274339-A 18 14-AUG-2001;		
FEATURES	Location/Qualifiers		
source	1..2625		
	/organism="unknown"		
	/mol_type="unassigned DNA"		
ORIGIN			
	Query Match	56.4%;	Score 2028.8; DB 6; Length 2625;
	Best Local Similarity	97.0%;	Pred. No. 0;
	Matches 2088;	Conservative 0;	Mismatches 62; Indels 2; Gaps 2;
Qy	15	GGCGGCAACTGAGCAAGGCTGAGGAGGAGCGCGGCGAGCGGAGCGGCTCGCGGCGAG	74
Db	234	GGAGGCGGAGCGCCCTGCGGTGGCGGCGCGGTGTCCGGCTCGCGCGCAGCGGCGCA	293
Qy	75	GAGCGCGCGGCGGCATGTGTCAACGCGGCTGCGTGGCAACCTTGGCACCGGCGACTGGT	134
Db	294	GGAATGTGACCGGCGCGTGTGTCAACGCGCGGTGCGTGCACCCCTGGGCACCGGCGAGTGGT	353
Qy	135	CTGCGCCCGCGGCTGGGCGAGCAATGCCAGCACTCGGGGGCGGCTTCAGACTAAC	194
Db	354	CTGCGCCCGCGGCTGGGCGGAGCAATGCCAGCACTCGGGGGCGGCTTCAGACTAAC	413
Qy	195	TGGATCTTCTGGGTTTGTGACAGATGGCACTGGAAATTAATAACAAGAGTGCAC	254
Db	414	TGGATCTTCTGGGTTTGTGACAGATGGCACTGGAAATTAATAACAAGAGTGCAC	473
Qy	255	GTGGCTCATTTAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGTCTAC	314
Db	474	GTGGCTCATTTAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGTCTAC	533
Qy	315	AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAAAATTTATGCAACCGCTAGT	374
Db	534	AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAAAATTTATGCAACCGCTAGT	593
Qy	375	TGCTGCATTTAGTGGCGCTCATTTGTTCTGAGAGATGGCAATGAGACTGTCCCTGAGGT	434
Db	594	TGCTGCATTTAGTGGCGCTCATTTGTTCTGAGAGATGGCAATGAGACTGTCCCTGAGGT	653
Qy	435	TGTTGCCACATCAGGTTATGCTTGTGCTATTTTATGATGGGACTCAAAATTTATGCAACCGCTAGT	494
Db	654	TGTTGCCACATCAGGTTATGCTTGTGCTATTTTATGATGGGACTCAAAATTTATGCAACCGCTAGT	713
Qy	495	TGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	554
Db	714	TGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	773
Qy	555	GTGTAAGATCAGTAATAGCAGGACTACTGTTTGAATGTGAATGTTCTGAAAACTGGAAGG	614
Db	774	GTGTAAGATCAGTAATAGCAGGACTACTGTTTGAATGTGAATGTTCTGAAAACTGGAAGG	833
Qy	615	TGAAGCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTTTCTTCATCGAGGACATCTG	674

Db	834	TGAAGCATGTGACATTCCTCACTGTACAGCAAACTGTGGTTTCTCTCATCGAGCATCTG	893
Qy	675	CAATTCAGTGTAGTGTAGAGGATGCTCTCTCTCTCAGACTGCGAGGTCCTCGATGTTTC	734
Db	894	CAATTCAGTGTAGTGTAGAGGATGCTCTCTCTCTCAGACTGCGAGGTCCTCGATGTTTC	953
Qy	735	AGTTCTCTGTACAGCTAACCACTCATTTTGGACTTCGAGAGGAATATTTCTAACTTAAAGCT	794
Db	954	AGTTCTCTGTACAGCTAACCACTCATTTTGGACTTCGAGAGGAATATTTCTAACTTAAAGCT	1013
Qy	795	CCCAGAGCATCTCATAAAGCTGTGCTCAATGGAACATTTATGTTGGTGTCTGAGGATA	854
Db	1014	CCCAGAGCATCTCATAAAGCTGTGCTCAATGGAACATTTATGTTGGTGTCTGAGGATA	1073
Qy	855	TATGTTCAACCACTCAGATTAATCACTGGTTCCTAGCGTATGACCTTGTCTTCTAGGAGTG	914
Db	1074	TATGTTCAACCACTCAGATTAATCACTGGTTCCTAGCGTATGACCTTGTCTTCTAGGAGTG	1133
Qy	915	GCTTCCCACTAAACCGTTCCTGTGAAACATGTGTTTGTAGATATGTCATTTTGGCATTT	974
Db	1134	GCTTCCCACTAAACCGTTCCTGTGAAACATGTGTTTGTAGATATGTCATTTTGGCATTT	1193
Qy	975	ATACAGGATAAAATTTTACATGATGAGGAAATTTGATTCACTGGGAATGTGACCAA	1034
Db	1194	ATACAGGATAAAATTTTACATGATGAGGAAATTTGATTCACTGGGAATGTGACCAA	1253
Qy	1035	TGAGTTGAGAGTTCCTTTCATATGATGATGATGATGATGATGATGATGATGATGATGATG	1094
Db	1254	TGAGTTGAGAGTTCCTTTCATATGATGATGATGATGATGATGATGATGATGATGATGATG	1313
Qy	1095	GGAGCAGTATGAGTGTGTTGGGCACTCTGACACATTTGTTACACTGGAAGATGCGCGAGT	1154
Db	1314	GGAGCAGTATGAGTGTGTTGGGCACTCTGACACATTTGTTACACTGGAAGATGCGCGAGT	1373
Qy	1155	GCTCATGCTGTGCTCATCTTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGACGA	1214
Db	1374	GCTCATGCTGTGCTCATCTTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGACGA	1433
Qy	1215	ATATGATTTGGATTAAGAACACATGAGTATATATACACACCCAGGTCCTTGTGCAAGG	1274
Db	1434	ATATGATTTGGATTAAGAACACATGAGTATATATACACACCCAGGTCCTTGTGCAAGG	1493
Qy	1275	GGTTTACGCGCATAGCAGTGTTCAGCACCATAGAGCCAGGCGCTTATACCTTCACTGCTG	1334
Db	1494	GGTTTACGCGCATAGCAGTGTTCAGCACCATAGAGCCAGGCGCTTATACCTTCACTGCTG	1553
Qy	1335	CTACAGGCTTTCAGTGCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT	1394
Db	1554	CTACAGGCTTTCAGTGCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT	1613
Qy	1395	GGATACCCAGATGTGGACCATTTTAAAGGACAGCGGATTTTCCGTTTCTTGCACACAGC	1454
Db	1614	GGATACCCAGATGTGGACCATTTTAAAGGACAGCGGATTTTCCGTTTCTTGCACACAGC	1673
Qy	1455	TGTGTAGTGTAGTGAACCATTCGCTGTTTGGGGGAAAACACACAAATGACATCTAT	1514
Db	1674	TGTGTAGTGTAGTGAACCATTCGCTGTTTGGGGGAAAACACACAAATGACATCTAT	1733
Qy	1515	GAGCCATCGGCCAAATGCTTCTCTTCAAGATTTTATGGCCCTATGACATTTGCTGTGACCG	1574
Db	1734	GAGCCATCGGCCAAATGCTTCTCTTCAAGATTTTATGGCCCTATGACATTTGCTGTGACCG	1793
Qy	1575	CTGGTCACTGTCTCCAGACCTGAT - TCCACCATGATGTCAACAGATTTGGCCATTCAG	1633
Db	1794	CTGGTCACTGTCTCCAGACCTGAT - TCCACCATGATGTCAACAGATTTGGCCATTCAG	1852
Qy	1634	CAGTCTTACACAGACCATGATGTGTTGGTGGTTCCTCAATAGTCTCTCTCTCTCAGCG	1693
Db	1853	CAGTCTTACACAGACCATGATGTGTTGGTGGTTCCTCAATAGTCTCTCTCTCTCAGCG	1912
Qy	1694	ACATCCTGTGATTTACCTTCGGAACAGTGTGATGCGCATTCGAGTGAAGCCGCTGTTTAG	1753



```
Db 1913 ACATCTGGTATTCACTCGGAACAGTGTGATGCGCATCGAGTGAAGCGCTGTTTAG 1972
Qy 1754 CAGCAGGACTGTGATTTCCGTGTGTCTGGAAACACAGGTCGCTCAGTGTATCTCGTGG 1813
Db 1973 CACAGGACCTGGTATTCCGTGTGTGTGGAAACACAGGTCGCTCAGTGTATCTCGTGG 2032
Qy 1814 CGTGGCAACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTCCAAAAGAACTC 1873
Db 2033 CGTGGCAACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTCCAAAAGAACTC 2092
Qy 1874 TTGACCATGACAGATGTGACCCAGCACACAGATGTTTACAGCTGCACAGCCACCAATG 1933
Db 2093 TTGACCATGACAGATGTGACCCAGCACACAGATGTTTACAGCTGCACAGCCACCAATG 2152
Qy 1934 ACTGCCACTGGTCAATGACCAATGTTGTCCTCCAGGAAACACAGCTGCTCAGAGGCCAGA 1993
Db 2153 ACTGCCACTGGTCAATGACCAATGTTGTCCTCCAGGAAACACAGCTGCTCAGAGGCCAGA 2212
Qy 1994 TCTCCATTTTGGTATGAGATGTCCTCCAGGAAACACAGCTGCTCAGAGGCCAGA 2053
Db 2213 TCTCCATTTTGGTATGAGATGTTGTCCTCCAGGAAACACAGCTGCTCAGAGGCCAGA 2272
Qy 2054 AGACCAAGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAGTGGAGCCCGGAATCAGG 2113
Db 2273 AGACCAAGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAGTGGAGCCCGGAATCAGG 2332
Qy 2114 AGTGCATTGCCCTGCCGAAATATCTGTGGCAATGCTGGCTGGCAATTTGGTTGG 2165
Db 2333 AGTGCATTGCCCTGCCGAAATATCTGTGGCAATGCTGGCTGGCAATTTGGTTGG 2384

RESULT 14
AR164813
LOCUS AR164813 2419 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 8 from patent US 6274339.
ACCESSION AR164813
VERSION AR164813.1 GI:16238022
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
AUTHORS Moore,K. and Nagle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
weight disorders, including obesity
JOURNAL Patent: US 6274339-A 8 14-AUG-2001;
FEATURES Location/Qualifiers
source 1..2419
/mol_type="unassigned DNA"

ORIGIN
Query Match 39.7%; Score 1428.8; DB 6; Length 2419;
Best Local Similarity 89.4%; Pred. No. 0;
Matches 1560; Conservative 0; Mismatches 182; Indels 2; Gaps 2;
Qy 189 ACTAATCGATCTCTCGGTTGTGACAGATGCACTGGAAATTAATAACAAACGAA 248
Db 677 ACTAATCGCTCTCTCGGTTGTGACAGATGCACTGGAAATTAATAACAAACGAA 736
Qy 249 GTGCACTGGCTCAATTAAGGACAGCCAAATAGAATTAATGAGACTCGTTCAATCATTT 308
Db 737 GTGCACTGGCTCAATTAAGGACAGCCAAATAGAATTAATGAGACTCGTTCAATCATTT 796
Qy 309 TGTACAGAGTGTAGTGGACCAATTTATGTTTATGATGGGACTCAATTTATGCAAC 368
Db 797 TGTACAGAGTGTAGTGGACCAATTTATGTTTATGATGGGACTCAATTTATGCAAC 856
Qy 369 GCTAGTGTGCTATTTAGTGGCTCATTTGCTCAGAGAGATGGCAATGAGACTGTGCC 428
Db 857 TCTGATTGTGCTTTAGTGGCTCATTTGCTCAGAGAGATGGCAATGAGACTGTGCC 916
Qy 429 TGAGGTTGTGACACATCAGTTATGCCCTTGCTGCAATTTTATTAGTATGCTGCTATAA 488
```

```
Db 917 TGAGTCACTGTCACTTCAGGTTATGCACTGCTGCAATTTTTCAGTGTATGCTTATAA 976
Qy 489 TTTGACTGGATTTAATATTACTTACAGTTTGTATGATGTTCCAAATTAATCTCAGGGCG 548
Db 977 TCTGACTGGATTTAATATTACTTACAAATTTTGACATGTGCTCCGAATAATTGCTCAGGCC 1036
Qy 549 AGGAGAGTGAAGATCAGTAATAGCAGCATCTGTTGAATGTGAATGTTCTGAAACTG 608
Db 1037 AGGAGAGTGAAGAGCAGTAACACAGCAGCGCTGTTGAGTGTGAATGTTCTGAAACTG 1096
Qy 609 GAAAGGTGAAGCATGTGACATTCCTCATCTGTAACAGCAACTGTGTTCTCTCATCGAGG 668
Db 1097 GAAAGGGGAGTGTGTGACATTCCTCATCTGTAACAGCAACTGTGTTCTCTCATCGAGG 1156
Qy 669 CATCTGCAATTTCAAGTGTGATGTGACAGGATGCTCTGCTTCTCAGACTGGCAGGTCCTGG 728
Db 1157 CATCTGTAATGCAAGCATACACAGAGGTGCTCTGCTTCTCCTCAGTGGCAGGTCCTGG 1216
Qy 729 ATGTTCAAGTCTCTGTACCCAGCTAACAGTCAATTTTGAATGCGAGAGAAATATTCTAATT 788
Db 1217 ATGTTCAATTCCTGTGCCAGCTAACAGTCTTTTGGACTCGAGAGAAATATTCTGATTT 1276
Qy 789 AAGCTCCCCAGAGCATCTCATAAAGCTGTGTTCAATGGAACATTAATGTGGGTGTTGG 848
Db 1277 AAGCTCCCCAGAGCATCTCATAAAGCTGTGTTCAATGGAACATTAATGTGGGTGTTGG 1336
Qy 849 AGGATATATGTTCAACCACTCAGATTAATTAACATGTTTCTAGCGTATGACCTTGTCTTAG 908
Db 1337 CGGATATATGTTCAACCATTCAGATTACAGCATGTTTATAGCGTATGACCTGACTTCTAG 1396
Qy 909 GGAGTGGCTTCCACTAAACCGTCTCTGTGAACATGTGGTTGTAGATATGCTCATCTTT 968
Db 1397 GGAAATGGCTTCCACTAAACCATTCCTGTGAACAGTGTGGTTGTGAAGATATGCTCATCTTT 1456
Qy 969 GGCATTATACAAAGATATAAATTTTACATGTATGGAGGAAATTTGATTCAACTGCGGAATGT 1028
Db 1457 GGCATTATACAAAGATATAAATTTTACATGTATGGAGGAAATTTGATTCAACAGGAACTGT 1516
Qy 1029 GACCAATGAGTTGAGAGTATTTTCAATTCATTAATGAGTCAATGGGTGTGTTGACCCCTAA 1088
Db 1517 GACCAATGAGTTGAGAGTATTTTCAATTCATTAATGAGTCAATGGGTGTGTTGAACTCCGAA 1576
Qy 1089 GGCAAGAGGACAGTATGCTGAGTGTGGGCTCTGACACATGTTTACATCTGAAAGATGG 1148
Db 1577 AGCTAAGATCAGTATGCTGAGTGTGGGCTCTGACACATGTTTACATCTGAAAGATGG 1636
Qy 1149 CCGAGTGTGCTGCTGCTCATCTTTGCTCATCTGCTCTCTATGGATATATAAGCAATGT 1208
Db 1637 CCGTGTGCTCATGTTGCTCATCTTTCGCTCATTTGCCACTCTATGGATATATAAGCGTTGT 1696
Qy 1209 GCAGGAATATGATTTGGATTAAGAACACATGAGTATATTACACCCAGGTCGCCCTTGT 1268
Db 1697 GCAGGAATATGATTTGGATTAAGAACACACATGAGTATATTACATCTCAGGTCGCTTGT 1756
Qy 1269 GCAAGGGGGTTACGGCCATAGCAGTGTATTACGACCATAGGACCGAGGCTCTATACGTTCA 1328
Db 1757 GCAAGGGGGTTATGCCCCACAGTAGTGTATTATGACAGGACCAAGGCTCTGACGTTCA 1816
Qy 1329 TGGTGGCTACAAGGCTTTTCAAGTGTGCAATTAAGTACCGGCTTGCAGATGATCTTACCGATA 1388
Db 1817 TGGTGGCTACAAGGCTTTTCAAGTGTGCAATTAAGTACCGGCTTGCAGATGATCTTACAGATA 1876
Qy 1389 TGATGTGGATACCCAGATGAGCACTTCTTAAGGACAGCGGATTTTCCGTTACTTGTCA 1448
Db 1877 CGATGTGGATCTCAGATGAGCACTTCTTAAGGACAGCGGATTTTCCGTTACTTGTCA 1936
Qy 1449 CACAGCTGTGATAGTGTGAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACAC 1508
Db 1937 TACAGCTGTGATAGTGTGAACCATGCTGGTGTGTTGGAGGGAACACACAAATGACAC 1996
Qy 1509 ATCTATGAGCATGGCCCAAAATGCTTCTCTCAGATTTTCATGGCTTATGACATGCTG 1568
```

```

Db      1997 TTCCATGAGCCAGGTGCCAAATGCTTCTCTCGACTTCATGGCTTATGACATTCGTTG 2056
Qy      1569 TGACCGTGTGTCAGTGTCTCCAGACCTGA-TTCCACCATGATGTCAAACAGATTGCGCA 1627
Db      2057 TGACCGATGTGTCAGTGTCTCCAGACCTGAGCTCCATCATGATGTCAAACAGATTGCGCA 2116
Qy      1628 TTCCAGCAGTCTTACAAACAGACACCATGATGTGTTCGGTGGTTTCAATGTCCTCC 1687
Db      2117 TT-CAGCAGTCTTGTAAACAGACCATGATGTGTTCGGCGCTTCAACAGCCTCCTCC 2175
Qy      1688 TCAGCAGCATCTGTTATTCACCTCGGAACAGTGTGATGGCATCGGAGTGAAGCGCTT 1747
Db      2176 TCAGTGAACGTCTTGGTCTTTTACCTCGGAGCAGTGGCATGCACACCGAGTGAAGCTGCTT 2235
Qy      1748 GTTTAGCAGCAGGACCTGTGTTATTCGGTGTGTGGAAACAGAGGTCTGCTCAGTGTATCT 1807
Db      2236 GTGTGGCAGCAGGACCTGTGTTATTCGGTGTGTGGGACACACAGTGTCTCGATGTACCT 2295
Qy      1808 CTGTGGCGCTGGCAACTGATGAACAGAGAAAGTTAAATCAGAAATGTTTTCAAA 1867
Db      2296 CCTGGAGTGGCAACTGAAGAACAGCAGAAAGTTAAATCAGAGTGTCTTTCAAA 2355
Qy      1868 GAACCTTGACCATGACAGATGTGACAGCAGACACAGATTGTTACAGCTGCACAGCCAA 1927
Db      2356 GAACCTTGACCATGACAGATGTGACAGCAGACACAGATTGTTACAGCTGCACAGCCAA 2415
Qy      1928 CCAA 1931
Db      2416 CCAA 2419

RESULT 15
AR164815
LOCUS      6370 bp      DNA      linear      PAT 17-OCT-2001
DEFINITION Sequence 12 from patent US 6274339.
ACCESSION AR164815
VERSION   AR164815.1 GI:16238024
KEYWORDS  Unknown.
SOURCE    Unknown.
ORGANISM  Unclassified.
REFERENCE 1 (bases 1 to 6370)
AUTHORS   Moore, K. and Nagle, D. Lynn.
TITLE     Methods and compositions for the diagnosis and treatment of body
          weight disorders, including obesity
JOURNAL   Patent: US 6274339-A 12-14-AUG-2001;
FEATURES  Location/Qualifiers
          source
            1..6370
              /organism="unknown"
              /mol_type="unassigned DNA"

ORIGIN

Query Match      37.98; Score 1364.4; DB 6; Length 6370;
Best Local Similarity 95.0%; Pred. No. 0;
Matches 1492; Conservative 7; Mismatches 48; Indels 23; Gaps 8;

Qy      2035 ATGTACTACTGTAAACAGAGACACAGCTGCAGGAGCTGTGCCCTGGACAGAACTGCCAG 2094
Db      1 ATGTACTACTGTAAACAGAGACACAGCTGCAGGAGCTGTGCCCTGGACAGAACTGCCAG 60
Qy      2095 TGGAGCCCGGAATCAGGATGTGATTCCTCCCGGAAATATCTGTGGCATTGGCTGG 2154
Db      61 TGGAGCCCGGAATCAGGATGTGATTCCTCCCGGAAATATCTGTGGCATTGGCTGG 120
Qy      2155 CATTGTGTGAAACTCATGTTTGAAATTTACTACTGCCAAGGAGAAATATGCAATGCT 2214
Db      121 CATTGTGTGAAACTCATGTTTGAAATTTACTACTGCCAAGGAGAAATATGCAATGCT 180
Qy      2215 AAATGTGTGTAGGAACCAATGCCCTTTGGCTTCTCTTACACCCAGAGAGAGGTA 2274
Db      181 AAATGTGTGTAGGAACCAATGCCCTTTGGCTTCTCTTACACCCAGAGAGAGGTA 240
Qy      2275 GAATTTGTCCTTAAGCAGCTGCGAATAATGAGTCACTCAGAGCATGTCCAGCTCACC 2334

```

```

Db      241 GAATTTGTCCTTAAGCAGCTGCGAATAATGAGTCACTCAGAGCATGTCCAAGCTCACC 300
Qy      2335 TTAACCCCATGGGTGCGCCTT---CGGAAGATCAATGTGTCTACTGG-----TGCTGGG 2386
Db      301 TTAACCCCATGGGTGCGCCTTCCGGAAGGTCAATGTGTCTACTKGCTGGGGAAG 360
Qy      2387 AAGATATGTGCCCATTTTACAAATAG-TTTACTACAGTGGATGCCGTCTG-----AGCCAG 2441
Db      361 GATATGTGCCCATTTTACAAATAGTTTACTACAGTGGATGSCCGTCTTGAGGGCCAG 420
Qy      2442 TGATGCT--GGATTTCTGTGAATTTT-----ATCAGAAACCCAGTACTCGGGGACTGAAGG 2494
Db      421 TGTGTCTTGTGATTTCTGTGGAAATTTTATTCAGGAACCCAGTTACTTCGGGGACTGAAGG 480
Qy      2495 CTGCAACCTGCAAT--CAACCCACTCAATGTGTAGTGTCTGTGAAAGGCTTCAAAACACAG 2552
Db      481 CTGCAACCTGCAATCAACCCACTCAATGTGTAGTGTCTGTGAAAGGCTTCAAAACACAG 540
Qy      2553 TGCTAA--GCAGTGGCGGACACCATGTGCTTGGAGCAGCATGTGGAGATTGCACACAGCG 2611
Db      541 TGCTAAGCAGTGGCGGACACCATGTGCTTGGAGCAGCATGTGGAGATTGCACACAGCG 600
Qy      2612 GCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTCCAATGCTTATG 2671
Db      601 GCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTCCAATGCTTATG 660
Qy      2672 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATAGATGAGCAGCCTGCCCTCG 2731
Db      661 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATAGATGAGCAGCCTGCCCTCG 720
Qy      2732 AAAATTGTTTCAAGCTACTGTACCTGTAGTCACTTGGAGCAACACAGGCTTGGCTGGT 2791
Db      721 AAAATTGTTTCAAGCTACTGTACCTGTAGTCACTTGGAGCAACACAGGCTTGGCTGGT 780
Qy      2792 GTACTGATCCAGCAATCTGCGCAAGGAAATGTCATAGAGGTTCTTATAAGGACCAG 2851
Db      781 GTACTGATCCAGCAATCTGCGCAAGGAAATGTCATAGAGGTTCTTATAAGGACCAG 840
Qy      2852 TGAAGATGCTTTCGCAAGCCCTACAGAAATTTCTATCCACAGCCCTGCTCAATTCGA 2911
Db      841 TGAAGATGCTTTCGCAAGCCCTACAGAAATTTCTATCCACAGCCCTGCTCAATTCGA 900
Qy      2912 GCATGTGTCTAGAGCAGCAGATACAATGCTGTTCTTCACTGCTCCAGCTTGCCTAAT 2971
Db      901 GCATGTGTCTAGAGCAGCAGATACAATGCTGTTCTTCACTGCTCCAGCTTGCCTAAT 960
Qy      2972 CCAAGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACCA 3031
Db      961 CCAAGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACCA 1020
Qy      3032 CAGCAAGCACTGCGAGACCTGCAATCTACGGTCTTCTACGGTGTATCCCAATGGAGGA 3091
Db      1021 CAGCAAGCACTGCGAGACCTGCAATCTACGGTCTTCTACGGTGTATCCCAATGGAGGA 1080
Qy      3092 AATGTACCAATGCAAGTGAATGGGACCGCTCTGTGCAACACCAACACACGCGGCAAGT 3151
Db      1081 AATGTACCAATGCAAGTGAATGGGACCGCTCTGTGCAACACCAACACACGCGGCAAGT 1140
Qy      3152 GCTTCTGCAACCAACGAGGCGCTCAAGGGGACGAGTGCAGCTATGTGAGTAGAAAAATC 3211
Db      1141 GCTTCTGCAACCAACGAGGCGCTCAAGGGGACGAGTGCAGCTATGTGAGTAGAAAAATC 1200
Qy      3212 GATACCAAGGAAACCTCTCAGAGGAACATGTTTATTATCTCTTCTTATTGACTATCAGT 3271
Db      1201 GATACCAAGGAAACCTCTCAGAGGAACATGTTTATTATCTCTTCTTATTGACTATCAGT 1260
Qy      3272 TCACCTTTTGTCTATCCCGAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGCTA 3331
Db      1261 TCACCTTTTGTCTATCCCGAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGCTA 1320
Qy      3332 TTCTCTGAGCAACAAAACAGGATTTGAGCATGTTTCATCAATGCTTCCAGAAATTTCAACC 3391

```



Db	1321	CTCCTCACGAACAAAACAGGGATTGGACATGTTTCATCAATGCTCCAAAGAAATTTCAACC	1380
Qy	3392	TCACATCACCTGGGCTGCCAGTTTCTCAGCTGRRACCCAGGCTGGAGAGATGCGCTG	3451
Db	1381	TCACATCACCTGGGCTGCCAGTTTCTCAGCTGRRACCCAGGCTGGAGAGATGCGCTG	1440
Qy	3452	TTGTTTCAAAAACCAACATTAAAGAGTACAAAGATAGTTTCTCTAATGAGAAAGTTTGATT	3511
Db	1441	TTGTTTCAAAAACCAACATTAAAGAGTACAAAGATAGTTTCTCTAATGAGAAAGTTTGATT	1500
Qy	3512	TTGCAACACCCCAATATCATCTTCTGTTTATGTCAGTAATTTTCACCTGGCCCATCA	3571
Db	1501	TTGCAACACCCCAATATCATCTTCTGTTTATGTCAGTAATTTTCACCTGGCCCATCA	1560
Qy	3572	AAATTCAGGT	3581
Db	1561	AAATTCAGAT	1570

Search completed: March 5, 2004, 05:11:01  
 Job time : 9065.63 secs

***This Page Blank (uspto)***



JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.  
FEATURES Location/Qualifiers  
source 1..3976  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..3976  
/gene="ATR1"  
/locus\_tag="HCM6607"  
ORIGIN  
Query Match 81.8%; Score 2942.2; DB 29; Length 3976;  
Best Local Similarity 84.4%; Pred. No. 0;  
Matches 2944; Conservative 0; Mismatches 545; Indels 0; Gaps 0;  
QY 93 TGTCAACGGCGTGGTGCACACCTGGCCACCGCCAGTGGTCTGCCCGCGGCTGGGT 152  
DB 1 TGTCAACGGCGTGGTGCACACCTGGCCACCGCCAGTGGTCTGCCCGCGGCTGGGT 60  
QY 153 GGGCGAGCAATGCGACACTGCGGGGGCGCTTCAGACTAACTGGATCTTCTGGGGTTGT 212  
DB 61 GGGCGAGCAATGCGACACTGCGGGGGCGCTTCAGNNNNNNNNNNNNNNNNNNNN 120  
QY 213 GACAGATGGACCTGGAAATATATAAATAAACAAGAGTGCACGTGGCTCATTTGAAGGACA 272  
DB 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180  
QY 273 GCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTCTACAGAGTGTGGAGACA 332  
DB 181 GCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTCTACAGAGTGTGGAGACA 240  
QY 333 TTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCAATTTAGTGGCT 392  
DB 241 TTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCAATTTAGNNNNNN 300  
QY 393 CATGTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTCACACATCAGTTA 452  
DB 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360  
QY 453 TGCCTTGTGCTATTTTGTAGTATGCTGCTATATAATTTGACTGGATTTAATTTACTTA 512  
DB 361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420  
QY 513 CAGTTTGTATGATGTGCCAAATAACTGTCTCAGCCGAGGAGAGTGTAAAGTCAAGTAATAG 572  
DB 421 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 480  
QY 573 CAGCGATCTGTTGAATGTGAATGTCTGAAATCTGAAAGAGTGAAGCATGTGACATTC 632  
DB 481 CAGCGATCTGTTGAATGTGAATGTCTGAAATCTGAAAGAGTGAAGCATGTGACATTC 540  
QY 633 TCACTGTACAGACAATGTGGTTTCTCATCGAGGCACTGCAATCAAGTGAATGTGAG 692  
DB 541 TCACTGTACAGACAATGTGGTTTCTCATCGAGGCACTGCAATCAAGTGAATGTGAG 600  
QY 693 AGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCGGATGTTTCAAGTTCCTGTACCACTAA 752  
DB 601 AGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCGGATGTTTCAAGTTCCTGTACCACTAA 660  
QY 753 CAGTCAATTTTGACTCGAGAGAAATTTCACTTAAGCTCCCGAGAGCATCTCATAA 812  
DB 661 CAGTCAATTTTGACTCGAGAGAAATTTCACTTAAGCTCCCGAGAGCATCTCATAA 720  
QY 813 AGCTGTGGTCAATGGAAACATTAATGTGGGTTGTGGAGGATATATGTTCAACCACTCAGA 872  
DB 721 AGCTGTGGTCAATGGAAACATTAATGTGGGTTGTGGAGGATATATGTTCAACCACTCAGA 780  
QY 873 TTATAACATGGTCTAGCGTATGACCTTCTAGGGAGTGGCTTCCACTAAACGGTTC 932  
DB 781 TTATAACATGGTCTAGCGTATGACCTTCTAGGGAGTGGCTTCCACTAAACGGTTC 840

QY 933 TGTGAACAATGTGGTGTGTAGATATGTCATTCTTTGGCATTATACAAGATATAATTTA 992  
DB 841 TGTGAACAATGTGGTGTGTAGATATGTCATTCTTTGGCATTATACAAGNNNNNNNNNN 900  
QY 993 CATGTATGGAGGAAAAATTCATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTC 1052  
DB 901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960  
QY 1053 CATTCATAATGATGATGGTGTGTGACCCCTPAGGCAAGAGCAGTATGCGATGGT 1112  
DB 961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020  
QY 1113 TGGGCACCTCTGCACACATTTGTACACTGAAGAATGGCCGAGTGGTCAATGCTCATCTT 1172  
DB 1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080  
QY 1173 TGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTGGATAGAA 1232  
DB 1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140  
QY 1233 CACATGGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTTTACGGCCATAGCAG 1292  
DB 1141 CACATGGAGTATATTACACACCCAGGGTCCCTTGTGCAAGGGGTTTACGGCCATAGCAG 1200  
QY 1293 TGTTTACGACCATAGGACCAAGGGCCCTATACGTTTCATGTGGCTACAAAGCTTTTCAGTGC 1352  
DB 1201 TGTTTACGACCATAGGACCAAGGGCCCTATACGTTTCATGTGGCTACAAAGCTTTTCAGTGC 1260  
QY 1353 CAATAAGTACGGCTTGCAGATGATCTTACCCATATGATGTGGATACCCAGATGTGGAC 1412  
DB 1261 CAATAAGTACGGCTTGCAGATGATCTTACCCATATGATGTGGATACCCAGATGTGGAC 1320  
QY 1413 CATTTCTTAAAGACACAGCCGATTTTTCGCTTACTTTCACACACAGCTGTGATAGTGGGAAC 1472  
DB 1321 CATTTCTTAAAGACACAGCCGATTTTTCGCTTACTTTCACACACAGCTGTGATAGTGGGAAC 1380  
QY 1473 CATCTGGTGTGGGGGAAACACACAAATGACATCTATGAGCCATGGCCCAATG 1532  
DB 1381 CATCTGGTGTGGGGGAAACACACAAATGACATCTATGAGCCATGGCCCAATG 1440  
QY 1533 CTTCCTCTCAGATTTTCATGGCTATGACATTTGCTGTGACCCGCTGTGCTGCCATGCCAG 1592  
DB 1441 CTTCCTCTCAGATTTTCATGGCTATGACATTTGNNNNNNNNNNNNNNNNNNNNNNNN 1500  
QY 1593 ACCTGATTCACCATGATGTCAACAGATTTGGCCATTCACAGCATCTTACACACAGCAC 1652  
DB 1501 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1560  
QY 1653 CATCTATGCTTGGTGGTTCATAGTCTCTCTCCTCAGCGACATCTCTGTATTTCACTTC 1712  
DB 1561 CATCTATGCTTGGTGGTTCATAGTCTCTCTCCTCAGCGACATCTCTGTATTTCACTTC 1620  
QY 1713 GGAACAGTGTGATCGCATCGGAGTGAAGCCGCTTGTAGCAGCAGGACCTGGTATTCG 1772  
DB 1621 GGAACAGTGTGATCGCATCGGAGTGAAGCCGCTTGTAGCAGCAGGACCTGGTATTCG 1680  
QY 1773 GTGTGTGTGAAACACAGGGTGGTCTCAGTGTATCTGTGGGCGCTGCACTCATGAACA 1832  
DB 1681 GTGTGTGTGAAACACAGGGTGGTCTCAGTGTATCTGTGGGCGCTGCACTCATGAACA 1740  
QY 1833 AGAAGAAAAGTAAAAATCAGAAATGTTTTTCCAAAAAGAACTCTTTGACCATGACAGATGTGA 1892  
DB 1741 AGAAGAAAAGTAAAAATCAGAAATGTTTTTCCAAAAAGAACTCTTTGACCATGACAGATGTGA 1800  
QY 1893 CCAGCACACAGATTTGTAGAGTGCACAGCAACCAATGACTGCGCCTGTGGCAATGA 1952  
DB 1801 CCAGCACACAGATTTGTAGAGTGCACAGCAACCAATGACTGCGCCTGTGGCAATGA 1860  
QY 1953 CCATTGTGTCCCAAGAAACACAGCTCTCAGAAAGCCAGATCTCCATTTTATAGGTATGA 2012  
DB 1861 CCATTGTGTCCCAAGAAACACAGCTCTCAGAAAGCCAGATCTCCATTTTATAGGTATGA 1920

```
QY 2013 GAATTGGCCCAAGGATAACCTTATGTACTACTGTAAACAAGAACACAGCTGCAGGAGTGT 2072
Db 1921 GAATTGGCCCAAGGATAACCTTATGTACTACTGTAAACAAGAACACAGCTGCAGGAGTGT 1980
QY 2073 TGCCTCTGGACACGAACTGCCAGTGGAGGCCCGGAAATCAGGAGTGCATTGCCCTGCCCGA 2132
Db 1981 TGCCTCTGGACACGAACTGCCAGTGGAGGCCCGGAAATCAGGAGTGCATTGCCCTGCCCGA 2040
QY 2133 AAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAACTCATGTTTGGAAATTTACTACTGC 2192
Db 2041 AAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAACTCATGTTTGGAAATTTACTACTGC 2100
QY 2193 CAAGGAGAAATTATGACAAATGCTAAATTTGTTCTGTAGCAACCAATGCCCTTTTGGCTTC 2252
Db 2101 CAAGGAGAAATTATGACAAATGCTAAATTTGTTCTGTAGCAACCAATGCCCTTTTGGCTTC 2160
QY 2253 TCTTACAACCCAGAGAGATGTCCCAATTTTGTCTTAAAGCAGCTGCCGAATTAATCAGTCAATC 2312
Db 2161 TCTTACAACCCAGAGAGATGTCCCAATTTTGTCTTAAAGCAGCTGCCGAATTAATCAGTCAATC 2220
QY 2313 TCAGAGCATGTCCAAAGCTCACCTTAACCCATGGTTCGGCTTCGGAAGATCAATGTGTC 2372
Db 2221 TCAGAGCATGTCCAAAGCTCACCTTAACCCATGGTTCGGCTTCGGAAGATCAATGTGTC 2280
QY 2373 CTACTGTGTCTGGGAAGATGTCCCAATTTTACAAATAGTTTACTACAGTGGATGCCGTC 2432
Db 2281 CTACTGTGTCTGGGAAGATGTCCCAATTTTACAAATAGTTTACTACAGTGGATGCCGTC 2340
QY 2433 TGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGAGCTGAA 2492
Db 2341 TGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGAGCTGAA 2400
QY 2493 GGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTGAAAGGCCCTGCAAAACACAG 2552
Db 2401 GGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTGAAAGGCCCTGCAAAACACAG 2460
QY 2553 TGCTAAGCAGTCCGGACACCATGTGCTTGAGGACAGCATGTGGAGATTGCAACGAGCG 2612
Db 2461 TGCTAAGCAGTCCGGACACCATGTGCTTGAGGACAGCATGTGGAGATTGCAACGAGCG 2520
QY 2613 CAGCTCTGAGTGCATGTGTGAGCAGCAACATGAGCAGTGTGTGACTCCAAATGCCATGT 2672
Db 2521 CAGCTCTGAGTGCATGTGTGAGCAGCAACATGAGCAGTGTGTGACTCCAAATGCCATGT 2580
QY 2673 GGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATAGATGAGCAGCCTGCCCCCTCA 2732
Db 2581 GGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATAGATGAGCAGCCTGCCCCCTCA 2640
QY 2733 AAATTGTTTCAAGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGGTG 2792
Db 2641 AAATTGTTTCAAGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGGTG 2700
QY 2793 TACTGATCCAGCAATCTGCGCAAGGAAATGCATAGAGGGTTTCTATAAAGGACCACT 2852
Db 2701 TACTGATCCAGCAATCTGCGCAAGGAAATGCATAGAGGGTTTCTATAAAGGACCACT 2760
QY 2853 GAAGATGCCCTTCGAGGCCCTTACAGAAATTTCTATCCAGCCCTGCTCAATTCAG 2912
Db 2761 GAAGATGCCCTTCGAGGCCCTTACAGAAATTTCTATCCAGCCCTGCTCAATTCAG 2820
QY 2913 CATGTGCTAGAGCAGCAGATACACTGCTTCTTCAATTCAGTCCAGCTTGGCAATG 2972
Db 2821 CATGTGCTAGAGCAGCAGATACACTGCTTCTTCAATTCAGTCCAGCTTGGCAATG 2880
QY 2973 CAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTGACCA 3032
Db 2881 CAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTGACCA 2940
QY 3033 AGCAAGCAGTCCGAGACCTGCATATCTGGCTTCTACGCTGATCCCAACCAATGGAGGAA 3092
Db 2941 AGCAAGCAGTCCGAGACCTGCATATCTGGCTTCTACGCTGATCCCAACCAATGGAGGAA 3000
QY 3093 ATGTACGCCATGCAAGTGCATAGGGCAGCGCTCTCTGTGTGCAACACCAACACGCGGCAAGTG 3152
```

```
Db 3001 ATGTGAGCCATGCAAGTGCATATGGCAGCGCTCTCTGTGCAACCAACACGCGCAAGTG 3060
QY 3153 CTTTGTGACCAACCAAGGCGCTCAAGGGGGAAGAGTGGCAGCTATGTGAGGTAGAAAATCG 3212
Db 3061 CTTTGTGACCAACCAAGGCGCTCAAGGGGGAAGAGTGGCAGCTATGTGAGGTAGAAAATCG 3120
QY 3213 ATACCAAGGAAACCTCTCAGAGGAACATGTTATTTATCTCTTCTTATGACTATCAGTT 3272
Db 3121 ATACCAAGGAAACCTCTCAGAGGAACATGTTATTTATCTCTTCTTATGACTATCAGTT 3180
QY 3273 CACCTTTAGTCTATCCAGGAAGATGCTGCTATTTACACAGCTATCAATTTTGGGTAC 3332
Db 3181 CACCTTTAGTCTATCCAGGAAGATGCTGCTATTTACACAGCTATCAATTTTGGGTAC 3240
QY 3333 TCCTGACGAAACAAACAGGGATTTGGACATGTTTCAATGCCCTCCAGAAATTTCAACCT 3392
Db 3241 TCCTGACGAAACAAACAGGGATTTGGACATGTTTCAATGCCCTCCAGAAATTTCAACCT 3300
QY 3393 CAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTCGAAGAGATGCCCTGT 3452
Db 3301 CAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTCGAAGAGATGCCCTGT 3360
QY 3453 TGTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3512
Db 3361 TGTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3420
QY 3513 TCGCAACCAACCAATATACCTTTCTTGTGTTATGTCAGTAATTTCACTGGCCCATCAA 3572
Db 3421 TCGCAACCAACCAATATACCTTTCTTGTGTTATGTCAGTAATTTCACTGGCCCATCAA 3480
QY 3573 AATTTCAGGT 3581
Db 3481 AATTTCAGAT 3489
```

```
RESULT 2
AY418588
LOCUS
DEFINITION
Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
AY418588
ACCESSION
AY418588
VERSION
AY418588.1 GI:39774548
SOURCE
GSS.
ORGANISM
Pan troglodytes (chimpanzee)
Pan troglodytes
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 3671)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
14671302
2 (bases 1 to 3671)
Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
source
1..3671
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
<1..>3671
gene
```

ORIGIN		/gene="ATRN" /locus_tag="HCM607"	
Query Match		69.7%; Score 2508.6; DB 29; Length 3671;	
Best Local Similarity		76.1%; Pred. No. 0;	
Matches 2517; Conservative		0; Mismatches 792; Indels 0; Gaps 0;	
Qy	273	GCCTAATAGATATATAGAGACTGCTTCAATCAATTTTGTACAGAGTGTAGTTGGACCA	332
Db	1	GCCTAATAGATATATAGAGACTGCTTCAATCAATTTTGTACAGAGTGTAGTTGGACCA	60
Qy	333	TTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCCT	392
Db	61	TTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCCT	120
Qy	393	CATTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTA	452
Db	121	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	180
Qy	453	TGCCCTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGATTTAATATTACTTA	512
Db	181	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	240
Qy	513	CAGTTTTCATATGTCCTCAATTAATCTGCTCAGCGCGAGGAGAGTGAAGATCAGTAATAG	572
Db	241	NNNTTTTGGTATGTCGGAATAACTGCTCAGCGCGAGGAGAGTGAAGATCAGTAATAG	300
Qy	573	CAGCGATCTGTTGAATGTAATGTTCTGAAACTGGAAGGTGAAGCATGTCACATTC	632
Db	301	CAGCGATCTGTTGAATGTAATGTTCTGAAACTGGAAGGTGAAGCATGTCACATTC	360
Qy	633	TCAGTGTACAGACAATCTGCTGTTCTCATCGAGGCATCTGCAATTCAGTGTCTAG	692
Db	361	TCAGTGTACAGACAATCTGCTGTTCTCATCGAGGCATCTGCAATTCAGTGTCTAG	420
Qy	693	AGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGATGTTTCAGTGTCTGTTACCACTAA	752
Db	421	AGGATGCTCTGCTTCTCAGACTGGCAGGNNCTGGATGTTTCAGTGTCTGTTACCACTAA	480
Qy	753	CCAGTCAATTTGGACTCGAGAGAAATCTTAATTAAGCTCCCGAGAGCATCTCATAA	812
Db	481	CCAGTCAATTTGGACTCGAGAGAAATCTTAATTAAGCTCCCGAGAGCATCTCATAA	540
Qy	813	AGCTGTGGTCAATGGAAACATTAATGTTGGTGTGTGGAGGATATATGTTCAACCACTCAG	872
Db	541	AGCTGTGGTCAATGGAAACATTAATGTTGGTGTGTGGAGGATATATGTTCAACCACTCAG	600
Qy	873	TTATAACATGGTTCTAGCTGATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTTC	932
Db	601	TTATAACATGGTTCTAGCTGATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTTC	660
Qy	933	TGTGAAACAATGTTGTTAGATATGTTCAATTTTGGCATTATACAAGGATAAAATTTA	992
Db	661	TGTGAAACAATGTTGTTAGATATGTTCAATTTTGGCATTATACAAGGNNNNNNNNNN	720
Qy	993	CATGTAATGGAGGAAAATTAATCACTGGGAATGTGACCAATGAGTGTGAGTGTTCAT	1052
Db	721	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	780
Qy	1053	CATTCAATAGTGTGCTGCTGTTGTTGACCCCTAAGGCAAGGAGCAGTATGCTGCTG	1112
Db	781	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	840
Qy	1113	TGGGCACTCTGCAACAATGTTTACCTGAAGAATGGCCGAGTGGTCACTGCTGCTATCTT	1172
Db	841	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	900
Qy	1173	TGGTCACTGCTCTCTATGATATATAAGCAATGTGCAAGGATATGATTTGGATAGAA	1232
Db	901	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	960
Qy	1233	CACATGGAGTATATTAACACCCAGGGTGGCTTGTGCAAGGGGGTTACGGCCATAGCAG	1292

Db	961	CACGTGGAGTATATTAACACCCAGGGTGGCTTGTGCAAGGGGGTTACGGCCATAGCAG	1020
Qy	1293	TGTTTACAGACCATAGGACCCAGGGCCCTATACGTTTCATGTTGGTGGCTACAAGGCTTTCAGTGC	1352
Db	1021	TGTTTACAGACCATAGGACCCAGGGCCCTATACGTTTCATGTTGGTGGCTACAAGGCTTTCAGTGC	1080
Qy	1353	CAATAAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGTGATACCCAGATGTGGAC	1412
Db	1081	CAATAAGTACCGGCTTCAGATGATCTCTACCGATATGATGTGTGATACCCAGATGTGGAC	1140
Qy	1413	CATTCTTAAAGACAGCCGATTTTCCGTTACTTCCACACAGCTGTGTATAGTGTGAAC	1472
Db	1141	CATTCTTAAAGACAGCCGATTTTCCGTTACTTCCACACAGCTGTGTATAGTGTGAAC	1200
Qy	1473	CATGCTGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCATGGCGCAATG	1532
Db	1201	CATGCTGTGTGTTGGGGGAAACACACAAATGACATCTATGAGCCATGGCGCAATG	1260
Qy	1533	CTTCTCTTCAGATTTTCATGGCTATGATGATTCGCTGTGACCGCTGTGTCAGTGTCTCCAG	1592
Db	1261	CTTCTCTTCAGATTTTCATGGCTATGATGATTCGCTGTGACCGCTGTGTCAGTGTCTCCAG	1320
Qy	1593	ACCTGATTCACCATGATGTCAAGATTTGGCCATTCAGAGCTCTTACACACAGC	1652
Db	1321	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	1380
Qy	1653	CATGTAATGTTTCCGTTGGTTTCAATAGTCTCTCTCAGGAGACATCTGTTGTTTACCTC	1712
Db	1381	CATGTAATGTTTCCGTTGGTTTCAATAGTCTCTCTCAGGAGACATCTGTTGTTTACCTC	1440
Qy	1713	GGAAACAGTGTGATCGCATCGGATGGAAGCGCTGTTTATGAGCAGGAGCTGGTATTCG	1772
Db	1441	GGAAACAGTGTGATCGGATGGAAGCGCTGTTTATGAGCAGGAGCTGGTATTCG	1500
Qy	1773	GTGTGTGTGGAAACACAGGCTGCTCAGTGTATCTGTTGGGCGCTGCAACTGATGAACA	1832
Db	1501	GTGTGTGTGGAAACACAGGCTGCTCAGTGTATCTGTTGGGCGCTGCAACTGATGAACA	1560
Qy	1833	AGAAAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTTTGACCATGACAGATGTGA	1892
Db	1561	AGAAAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTTTGACCATGACAGATGTGA	1620
Qy	1893	CCAGCACAACATTTGTACAGTGCACAGCCCAACCACTGACCTGCCACTGGTGCATGA	1952
Db	1621	CCAGCACAACATTTGTACAGTGCACAGCCCAACCACTGACCTGGTGCATGA	1680
Qy	1953	CCATTGTCTCCCAAGAACACACAGCTGCTCAGAAAGGCGCAGATCTCCATTTTATGATGA	2012
Db	1681	CCATTGTCTCCCAAGAACACACAGCTGCTCAGAAAGGCGCAGATCTCCATTTTATGATGA	1740
Qy	2013	GAATTTGCCCAAGGATAACCTTATGTACTGTAAACAAGNAGACACAGCTGCAGAGCTG	2072
Db	1741	GAATTTGCCCAAGGATAACCTTATGTACTGTAAACAAGNAGACACAGCTGCAGAGCTG	1800
Qy	2073	TGCCCTGACACAGAACTGCGAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCCGA	2132
Db	1801	TGCNNNGAACAGAACTGCGAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCCGA	1860
Qy	2133	AAATATCTGTGGCATTTGGCTGGCAATTTGGTTGGAAACTCATGTTTGAATAATCTACTGC	2192
Db	1861	AAATATCTGTGGCATTTGGCTGGCAATTTGGTTGGAAACTCATGTTTGAATAATCTACTGC	1920
Qy	2193	CAAGGAGATTTATGACAAATGCTAAATTTCTGTAGGAAACCACAATGCCCTTTGGGCTTC	2252
Db	1921	CAAGGAGATTTATGACAAATGCTAAATTTCTGTAGGAAACCACAATGCCCTTTGGGCTTC	1980
Qy	2253	TCCTTACAAACCCAGAGAGGTAGTAATTTGCTTAAAGCAGCTGGGAATAATGCAAGTCACT	2312
Db	1981	TCCTTACAAACCCAGAGAGGTAGTAATTTGCTTAAAGCAGCTGGGAATAATGCAAGTCACT	2040
Qy	2313	TCAGAGCATGTCCAGCTCACCTTAAACCCCATGGTGGCTTCGGGAAGATCAATGTGTC	2372

```
Db 2041 TCNNNNNNNTCAAGCTCACTTAACCCCAATGGGTCGGCCCTTCGGGAAGATCAATGTGTGC 2100
Qy 2373 CTACTGGTGTGGGAAGATATGTCCCACTTTACAAATAGTTTACTACAGTGTGATCCGTC 2432
Db 2101 CTACTGGTGTGGGAAGATATGTCTCCATTTACAAATAGTTTACTACAGTGTGATCCGTC 2160
Qy 2433 TGAGCCCAAGTATGTCTGGATTTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2492
Db 2161 TGAGCCCAAGTATGTCTGGATTTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAA 2220
Qy 2493 GGCTGCAACCTGCATCAACCCCACTCAATGTGTAGTGTCTGTAAGAGCCCTGCACCAACAG 2552
Db 2221 GGCTGCAACCTGCATCAACCCCACTCAATGTGTAGTGTCTGTAAGAGCCCTGCACCAACAG 2280
Qy 2553 TGCTAAGAGTGCAGGACACCAATGTGCTTGTAGGACAGCATGTGAGATGTGCACGAGCG 2612
Db 2281 TGCTAAGAGTGCAGGACACCAATGTGCTTGTAGGACAGCATGTGAGATGTGCACGAGCG 2340
Qy 2613 CAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAGTGTGTGGACTCCAAATGCCATGT 2672
Db 2341 CAGCTCTGAGTGCATGTGTGAGCAACATGAAGCAGTGTGTGGACTCCAAATGCCATGT 2400
Qy 2673 GGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCTGAA 2732
Db 2401 GGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCTGAA 2460
Qy 2733 AATTGTTTCAAGCTACTGTACCTGTAGTCTGTTGGAGCAACAGGCTGTGGCTGGTG 2792
Db 2461 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2520
Qy 2793 TACTGATCCAGCAATCTGCGCAAGGAAATGCATAGAGGGTTCTCTATAAGAGGACCACT 2852
Db 2521 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2580
Qy 2853 GAAGATGCTCTGCAAGCCCTTACAGAAATTTCTATCCACAGCCCTCTCAATTCAG 2912
Db 2581 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2640
Qy 2913 CATGTGCTGTAGAGCAGCAGATACAACTGGTCTTTCTTCACTGTCCAGCTTGGCAATG 2972
Db 2641 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2700
Qy 2973 CAACGCCCAAGTAAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTGACCA 3032
Db 2701 CAACGCCCAAGTAAATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTGACCA 2760
Qy 3033 AGCAGGACCTGCAGACCTGCATATCTGCTTCTACGCTGATCCACCAATGGAGGA 3092
Db 2761 AGCAGGACCTGCAGACCTGCATATCTGCTTCTACGCTGATCCACCAATGGAGGA 2820
Qy 3093 ATGTCAGCCATGCAAGTGAATGGCAGCGCTCTCTGTGCAACACCAACACGCGCAAGTG 3152
Db 2821 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2880
Qy 3153 CTCTGTGCAACCAAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGGTGAAATCG 3212
Db 2881 CTCTGTGCAACCAAGGCGTCAANNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 2940
Qy 3213 ATACCAAGGAAACCTCTCAGAGGAACATGTTATATCTCTCTTTATTGACTATCAGTT 3272
Db 2941 ATACCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 3000
Qy 3273 CACTTTTGTATCTCCAGGAAGATGATGCTATTATACAGCTATCAATTTTGTGCTAC 3332
Db 3001 CACTTTTGTATCTCCAGGAAGATGATGCTATTATACAGCTATTAATTTTGTGCTAC 3060
Qy 3333 TCCTGAGCAACAAACAGGGATTTGGACATGTTTCATCAATGCTCCCAAGAAATTCACCT 3392
Db 3061 TCCTGAGCAACAAACAGGGATTTGGACATGTTTCATCAATGCTCCCAAGAAATTCACCT 3120
Qy 3393 CAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGT 3452
Db 3121 CAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGT 3180
```

```
Qy 3453 TGTTTCAAAACCAACATTAAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTGATTT 3512
Db 3181 TGTTTCAAAACCAACATTAAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTGATTT 3240
Qy 3513 TCSCAACCCACCAATATCACTTCTTGTGTTATGTCTAGTAATTTCACTGGCCCAATCAA 3572
Db 3241 TCSCAACCCACCAATATCACTTCTTGTGTTATGTCTAGTAATTTCACTGGCCCAATCAA 3300
Qy 3573 AATTCAGGT 3581
Db 3301 AATTCAGAT 3309

RESULT 3
AY418589
LOCUS
DEFINITION Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
genomic survey sequence.
ACCESSION AY418589
VERSION AY418589.1 GI:39774549
SOURCE GSS.
KEYWORDS Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3976)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 3976)
AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT These sequences were made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source
1..3976
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>3976
/gene="ATRN"
/locus_tag="HCM6607"

gene
Query Match 69.7%; Score 2507; DB 29; Length 3976;
Best Local Similarity 76.6%; Pred. No. 0;
Matches 2672; Conservative 0; Mismatches 817; Indels 0; Gaps 0;

Qy 93 TGTCAACGGCGGTGCTGCAACCTGGCACCGGCAGTGCCTGTCGCCCGCGCTGGGT 152
Db 1 TGTCAACGGCGGTGCTGCAACCTGGCACCGGCAGTGCCTGTCGCCCGCGCTGGGT 60
Qy 153 GGGGAGCAATGCCAGCACTGCGGGGCGCTTCAGACTAACTGATCTTTCTGGGTTGT 212
Db 61 GGGGAGCAATGCCAGCACTGCGGGGCGCTTCAGNNNNNNNNNNNNNNNNNNNNNNNN 120
Qy 213 GACAGATGCACTGGAAATTATAATACAAACGAAGTGCAGTGGCTCATTTGAAGACA 272
Db 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180
Qy 273 GCCAAATAGAATAATAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTGGACCA 332
Db 1
```



181 GCCAAATAGATAATGAGACTTCGCTCAACCAATTTTGCTACAGAAATGAGCTGGGACCA 240  
333 TTTATATGTTATGATGGGACTCAATTTATGCAACCGCTAGTTGCTGCAATTTAGTGGCT 392  
241 TTTATATGTTATGATGGGACTCAATTTACGCACTCTGATGCTGCTTTAGNNNNNN 300  
393 CATTTGTTCTGAGAGATGCAATGAGACTGTCCTCGAGGTTGTTGGCACATCAGGTTA 452  
301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360  
453 TGCCTGCTGCAATTTTATGATGCTGCTTATAATTTGACTGGAATTAATATTACTTA 512  
361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420  
513 CAGTTTGTATGATGTCCTCAATTAACCTGCTCAGGCGGAGAGAGTGAAGTCAAGTAAAG 572  
421 NNNNTTTGACATGTCGGAATAATTTGCTCAGGCCGAGGAGAGTGAAGAGCAAGTAAACAG 480  
573 CAGCGATACCTGTTGAATGTCGAATGTTCTGAAAACTGGAAGGTGAAGCATGTGACATTTCC 632  
481 CAGCAGCGCTGTTGAGTGTGAATGTTCTGAAACTGGAAGGGGAGTCTGTTGACATTTCC 540  
633 TCACTGACAGACAACCTGCTGCTTTCTCTATCGAGGCACTCTGCAATTCAGTGAATGTAG 692  
541 TCACTGACAGACAACCTGCTGCTTTCTCTACCGAGGCACTCTGTAATGCAAGCGATACCAAG 600  
693 AGGATGCTCCTGCTTCTCAGACTGGCAGGGTCTCTGGATGTTTCACTTCTCTGACAGCTAA 752  
601 AGGTGCTCCTGCTTCTCCTCTCAGTGGCAGGGTCTCTGATGTCATTTCTCTGCTGACCTAA 660  
753 CCACTCATTTTGGACTCGAGAGAAATATTTCTAACTTAAAGTCTCCCGAGAGCACTCATAA 812  
661 CCACTCTTTTGGACTCGAGAGAAATATTTCTGATTTAAAGCTTCCCGAGAGCTCTCATAA 720  
813 AGCTGTGCTCAATGGAAACATATGTTGGTGTGTTGGAGGATATGTTCAACCACTCAGA 872  
721 AGCTGTGCTCAATGGAAACATATGTTGGTGTGTTGGCGGATATGTTCAACCACTCAGA 780  
873 TTATAACATGTTCTAGCTATGACCTTGTCTTCTAGGGAGTGGCTTCCATAACCGTTTC 932  
781 TTACAGCATGGTTTACGCTATGACCTGACTTCTAGGGAATGGCTTCCATAACCACTTC 840  
933 TGTGAACAATGTTGTTGATATGTTGATATGTTGCTATTTCTTGGCAATATACAAGATATA 992  
841 TGTGAAACAGTGTGTTGATATGTTGATATGTTGCTATTTCTTGGCAATATACAAGNNNNNN 900  
993 CATGTATGGAGGAAAAATTCATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCA 1052  
901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960  
1053 CATTCATAATGAGTCACTGGGTGTTGTTGACCCCTAAGGCAAGGAGCATATGCACTGGT 1112  
961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020  
1113 TGGSCACTCTGCAACATTTGTTACACTGAAGAATGGCCGAGTGGTCACTGCTGCTATCTT 1172  
1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080  
1173 TGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAAGGAATATGATTTGGATAGAA 1232  
1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140  
1233 CACATGAGTATATACACACCAAGGTCCTTGTGCAAGGGGTTACGGCCATAGCAG 1292  
1141 CACATGAGTATATATCACTCAGGGTGTCTGTGCAAGGGGTTATGCCCACAGTAG 1200  
1293 TGTTTACGACCATAGGACCGGCTTATCGTTTCACTGTTGGCTACAAGGCTTTCACTGC 1352  
1201 TGTTTATGATCAACAGGACCAAGGCTCTGATCTTCACTGTTGGCTACAAGGCTTTCAAGCG 1260  
1353 CAATAGTACCGGCTTGCAGATGCTCTACCGATATGATGTTGATACCAAGATGTTGAC 1412  
1261 CAACAAATACCGGCTTGCAGATGACCTCTACAGATAGTGTGATCTCAGATGTTGAC 1320

1413 CATTTTAAAGACACCGGATTTTTCGTTACTTTCGACACACAGCTGATAGTGTGGAAC 1472  
1321 CATTTTAAAGACACCGGATTTTTCGTTACTTTCGATACAGCTGTGATAGTGTGGAAC 1380  
1473 CATGCTGTGTTTGGGGGAAACACACACAATATGACACATCTATGAGCCATGGCCAAATG 1532  
1381 CATGCTGTGTTTGGAGGAAACACACACAATATGACACATCTTCCATGAGCCACGGTGC 1440  
1533 CTTCTCTTCAATTTTCACTGCGCTATGACATTCCTCTGTCGCGCTGCTGCTGCTTCCAG 1592  
1441 CTTCTCTTCACTTCACTGCGCTTATGACATTTGNNNNNNNNNNNNNNNNNNNNNNNNNN 1500  
1593 ACCTGATTTCCACCATGATGTCAACAGATTTTGGCCATTTCCAGCAGCTTTACACAACAG 1652  
1501 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1560  
1653 CATGATGTTGTTTCGGTGGTTTCAATAGTCTCTCTCAGCGACATCTCTGTTATTCAGCTC 1712  
1561 CATGATGTTGTTTCGGCGGCTTCAACAGCCTCTCTCTCAGTGACGCTTTGGTCTTTACCTC 1620  
1713 GGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTAGCAGCAGCACTGCTATTCG 1772  
1621 GGAGCAGTGCATGACACACCGCAGTGAAGTGTGTTGTTGCGCAGCAGCACTGCTATCCG 1680  
1773 GTGTGTGTGGAACACACAGGCTGCTCTCAGTGTATCTCTGCGCGCTGCGCACTGATGAACA 1832  
1681 GTGTGTGTGGAACACACAGTGTCTCTGATGTACCTCTCTGGAGTTGGCACTGAAGAACA 1740  
1833 AGAAGAAAGTTAAATCAGATGTTTTCCTCAAGAACTCTTGACCATGACAGATGTA 1892  
1741 AGCAGAAAGTTAAATCAGAGTGTCTTCTAAAGAACTCTTGACCATGACAGATGTA 1800  
1893 CCAGCACACAGATTTTACAGCTGACACAGCAACCAATGACCTGCACTGCTGCAATGA 1952  
1801 CCAGCACACAGATTTTACAGCTGACACAGCAATACCAATGACCTGCACTGCTGCAATGA 1860  
1953 CCATGTTGCTCCAGGAAACACAGCTGCTGAGAGGCGAGATCTTCAATTTTAGGTATGA 2012  
1861 TCCTGTGTCTCTGTAACCAAGCTGACAGAGGCGAGATCTTCCATTTGCCAAGTATGA 1920  
2013 GAATTTGCCCAAGGATAACCTTACTGTACTGTAAACAGAAACAGCTGCGAGGAGCTG 2072  
1921 GAGTTGCCCAAGGATTAACCCATGCTACTGTCAATGABAAACAGCTGCGAGGAGCTG 1980  
2073 TGCCCTGGACACAGAACTGCGAGTGGAGCCCGGAATCAGAGTGCATTTGCCCTGCCCGA 2132  
1981 TGCCCTAGACCAAGCTGCGAGTGGAGCCCGGAATCAAGAGTGCATGCGCCCTGCCGGA 2040  
2133 AAATATCTGTGCGCATTCGCTGCGCATTTTGGTTGGAACTCTATGTTTGAATAATCTACTGC 2192  
2041 AAATATCTGTGCGCAATGCTGGCATTTTGGTTGGAACTCTGTGCTGAAATCACTACTGC 2100  
2193 CAAGGAGATTTATGCAATGCTAAATTTGTTCTGTAGGAACCAATGCCCTTTGGCTTC 2252  
2101 TAAGGAGATTTATGCAATGCTAAATTTGCTGTAGGAACCAATGCCCTTTGGCTTC 2160  
2253 TCTTACAAACCAAGAGGATGATTTGCTTAAAGCAGCTGCAATTAATGAGTGCATC 2312  
2161 CCTCACATCCCAGAGAGGAGTGTGCTCTTAAAGCAGCTTCAATTAATGCAATCATC 2220  
2313 TCAGAGCATGTTCAAGCTCACTTTAAACCCATGCGCTGCGCTTCGGAAGATCAATGTGC 2372  
2221 TCAAAGTATGTTCAAGCTCACTCTGACTCTGATGGTTGGTCTTCGGAAGATCAATGTGC 2280  
2373 CTACTGTGCTGGAGATGATGTCCTTCAATTAAGTATTACTACAGTGGATGCCGTC 2432  
2281 TTACTGTGCTGGAGGATGATGTCCTTCAATTCACAAATAGTGTGCTGAGTGGATGCCATC 2340  
2433 TGAGCCCAAGTGTGATTTCTGTGGAATTTTATCAGAAACCCAGTACTCGGGAGCTGAA 2492  
2341 TGAGCCCAAGTGTGATTTCTGTGGATCTTGTGAGGATCTTGTGAGGCTTAGTACTCGGGATTA 2400

```
Qy 2493 GGTGCAACCTGCATCAACCCATCAATAGTAGTGTGTGTGAAGGCGCTGCAAAACACAG 2552
Db 2401 GGTGCAACCTGCATCAACCCCTCAATGGCAGCGTCTGTGAAGGCGCTGCAAAACACAG 2460
Qy 2553 TGTAAAGCAGTGGCGACACCATGTGCTTGAGCAGCATGTGGAGATTGACACAGCGG 2612
Db 2461 TGCACAGCAGTGGCGACACCATGTGCTTGAGCAGCATGTGGAGATTGACACAGCGG 2520
Qy 2613 CAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGAGATCCAAATGCCATGT 2672
Db 2521 CAGCTGGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGAGATCCAAATGCCATGT 2580
Qy 2673 GGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCATGCCCCCTCA 2732
Db 2581 GGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCATGCCCCCTCA 2640
Qy 2733 AAATGTTTCAGGCTACTGTACCTGTAGTCAATGCTTTGGAGCAACACAGGCTGTGGTGTG 2792
Db 2641 AAATGTTTCAGGCTACTGTACCTGTAGTCAATGCTTTGGAGCAACACAGGCTGTGGTGTG 2700
Qy 2793 TACTGATCCAGCAATACTGGCAAGGGAATGATAGAGGTTCTATTAAGGACCAT 2852
Db 2701 TACTGATCCAGCAATACTGGCAAGGGAATGATAGAGGTTCTATTAAGGACCAT 2760
Qy 2853 GAAGATGCCCTTCGCAAGCCCTACAGAAATTTCTATCCACAGCCCTGCTCAATGCCAG 2912
Db 2761 GAAGATGCCCTTCGCAAGCCCTTCGCAAGGAAATGTTATCCACAGCCCTTCTGACTCCAG 2820
Qy 2913 CATGTGCTTAGAGGACAGCAGATACAACTGGTCTTTTCATTCATCTGTCAGCTGCGAATG 2972
Db 2821 CATGTGCTTAGAGGACAGCAGATACAACTGGTCTTTTCATTCATCTGTCAGCTGCGAATG 2880
Qy 2973 CAAGCGCCAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGAGAC 3032
Db 2881 CAAGCGCCAGTAAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGAGAC 2940
Qy 3033 AGGCAAGCACTGGGAGCCTGCATATCTGCTTCTACGGTGTATCCACCAATGGAGGAA 3092
Db 2941 AGGCAAGCACTGGGAGCCTGCATATCTGCTTCTACGGTGTATCCACCAATGGAGGAA 3000
Qy 3093 ATGTCAGCCATGAAGTGCANTGGCAGCGTCTCTGTGCAACACCAACACGCGGAGTG 3152
Db 3001 ATGTCAGCCATGAAGTGCANTGGCAGCGTCTCTGTGCAACACCAACACGCGGAGTG 3060
Qy 3153 CTTCTGCACCAACCAAGGCGTCAAGGGGACGAGTCCAGCTATGTGAGTAGAATAATCG 3212
Db 3061 CTTCTGTACCAACCAAGGCTCAAGGGGACGAGTCCAGCTATGTGAGTAGAATAATCG 3120
Qy 3213 ATACCAAGGAAACCTCTCAGAGGACATGTTATTTATATCTCTCTTATTGACTATCAGTT 3272
Db 3121 ATACCAAGGAAACCTCTCAGAGGACATGTTATTTATATCTCTCTTATTGACTATCAGTT 3180
Qy 3273 CACCTTTAGTCTATCCAGGAGATGATCGCTATTTACAGAGTATCAATTTTGTGGCTAC 3332
Db 3181 CACCTTTAGTCTATCCAGGAGACGACCGCTACTACAGAGCATCACTTTGTGGCTAC 3240
Qy 3333 TCCTGACGAAACAAACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTCACCT 3392
Db 3241 TCCTGATGAACAAACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTCACCT 3300
Qy 3393 CACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTCGAGAGAGATGCTCTGT 3452
Db 3301 CACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTCGAGAGAGATGCTCTGT 3360
Qy 3453 TGTTTCAAAACCAACATTAAGGAGTACAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3512
Db 3361 TGTTTCAAAACCAACATTAAGGAGTACAAGATAGTTTCTCTAATGAGAAGTTTGATTT 3420
Qy 3513 TCGCAACCAACCAAAATATCATCTTTCTTTGTTTATGTCAGTAATTCACCTGGCCCATCAA 3572
Db 3421 TCGCAACCAACCAAAATATCATCTTTCTTTGTTTATGTCAGTAATTCACCTGGCCCATCAA 3480
Qy 3573 AATTCAGGT 3581
```

```
Db 3481 AATTCAGAT 3489
|||||
RESULT 4
BX440935
LOCUS
DEFINITION
BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone
CSODF012YL19 5-PRIME, mRNA sequence.
ACCESSION
BX440935
VERSION
BX440935.1 GI:30781889
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1201)
Li.W.B., Gruber.C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished (2001)
CONTACT: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10212.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSODF012CF10QPI&cluster=10212.f. Contact :
Feng liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSODF012CF10QPI.
FEATURES
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODF012YL19"
/tissue_type="FETAL BRAIN"
/dev_stage="fetal"
/clone_lib="Homo sapiens FETAL BRAIN"
/note="Organ: brain; Vector: pcwSPORT_6; 1st strand cDNA
was primed with a NotI-oligo (dT) primer. Five prime end
enriched, double-strand cDNA was digested with Not I and
cloned into the Not I and EcoRV sites of the pcwSPORT 6.
vector. Library was not normalized."
ORIGIN
Query Match 25.7%; Score 925.6; DB 13; Length 1201;
Best Local Similarity 98.4%; Pred. No. 4.7e-223;
Matches 962; Conservative 4; Mismatches 8; Indels 4; Gaps 3;
Qy 2115 GTGCATTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGAACCTCATG 2174
Db 57 GTGCATTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTGAACCTCATG 116
Qy 2175 TTGTAAATTAATCTACTGTCACAGGAGATTTATGCAATGCTAAATTTCTCTAGAACCA 2234
Db 117 TTGTAAATTAATCTACTGTCACAGGAGATTTATGCAATGCTAAATTTCTCTAGAACCA 176
Qy 2235 CAATGCCCTTTGGCTTCTCTTAAACCCAGAGGTAGATAATTTGTCCTTAAGCAGCT 2294
Db 177 CAATGCCCTTTGGCTTCTCTTAAACCCCTGATGAAGTAGAATTTGTCCTTAAGCAGCT 236
Qy 2295 GCGAATAATGCAATCTCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGTGGGCT 2354
Db 237 GCGAATAATGCAATCTCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGTGGGCT 296
Qy 2355 TCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGCCCCATTAACAATAGTTT 2414
Db 297 TCGGAAGATCAATGTGCTCTACTGCTGGGAAGATATGCCCCATTAACAATAGTTT 356
Qy 2415 ACTCAGTGGATGCCGCTCTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAAC 2474
```

Db 357 ACTACATGATCCGCTCTGAGCCAGTGTGATCTGTGGATTTTATCAGAAC 416  
Qy 2475 CAGTACTCGGGAGTGAAGCTGCAACCTGATCAACCCACTCAATGGTAGTCTGTGA 2534  
Db 417 CAGTACTCGGGAGTGAAGCTGCAACCTGATCAACCCACTCAATGGTAGTCTGTGA 476  
Qy 2535 AAGGCTCGAAACACACAGTCTTAAGCAGTCCGAGACACCATGCTTGGAGACACATG 2594  
Db 477 AAGGCTCGAAACACACAGTCTTAAGCAGTCCGAGACACCATGCTTGGAGACACATG 536  
Qy 2595 TGGAGATTGACACAGCGAGCTCTGTAGTGTGATGTGTGAGCAACATGAAGCAGTGTGT 2654  
Db 537 TGGAGATTGACACAGCGAGCTCTGTAGTGTGATGTGTGAGCAACATGAAGCAGTGTGT 596  
Qy 2655 GGACTCCAATGCCTATGTGGCTCTTCCCTTTGGCCAGTGTGGAATGATACGAT 2714  
Db 597 GGACTCCAATGCCTATGTGGCTCTTCCCTTTGGCCAGTGTGGAATGATACGAT 656  
Qy 2715 GAGCAGCTGCCCCCTGAAATGTTTCCAGCTACTGTACCTGTAGTCAATGCTTGGAGCA 2774  
Db 657 GAGCAGCTGCCCCCTGAAATGTTTCCAGCTACTGTACCTGTAGTCAATGCTTGGAGCA 716  
Qy 2775 ACAGGCTGTGGTGTGTACTGATCCAGCAGTATCTGGCAAGGGAATGATAGAGG 2834  
Db 717 ACCAGGCTGTGGTGTGTACTGATCCAGCAGTATCTGGCAAGGGAATGATAGAGG 776  
Qy 2835 TTCTATTAAGGACCACTGAGATGCTTCCAGGAGCTTCCAGGAAATTTCTATCCACA 2894  
Db 777 TTCTATTAAGGACCACTGAGATGCTTCCAGGAGCTTCCAGGAAATTTCTATCCACA 836  
Qy 2895 GCGCTGCTCAATTCAGCATGTGTGTAGAGCAGCAGATCAATGCTGTCTTCAATCA 2954  
Db 837 GCGCTGCTCAATTCAGCATGTGTGTAGAGCAGCAGATCAATGCTGTCTTCAATCA 896  
Qy 2955 CTGTCCAGCTGCAATGCAAGCCACAGTAAATGATCAATCAATGAGCATCTGTGAGAA 3014  
Db 897 CTGTCCAGCTGCAATGCAAGCCACAGTAAATGATCAATCAATGAGCATCTGTGAGAA 956  
Qy 3015 GTGTGAGAACTGACACAGGCAAGCACTCGAGAGCTGATATCTTGGCTTCTACGGTGA 3074  
Db 957 GTGTGAGAACTGACACAGGCAAGCACTCGAGAGCTGATATCTTGGCTTCTACGGTGA 1013  
Qy 3075 TCCACCAATGGAGGAA 3092  
Db 1014 T-CCACCAATGGAGGAA 1030

RESULT 5  
LOCUS BU506373  
DEFINITION AGENCOURT\_10015485 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:849441  
5', mRNA sequence.  
ACCESSION BU506373  
VERSION BU506373.1 GI:22812606  
KEYWORDS EST.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 964)  
NIH-MGC http://mgi.nci.nih.gov/  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabs-remail.nih.gov  
Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov

plate: L14M14050 row: k column: 18  
High quality sequence stop: 592.  
FEATURES  
source Location/Qualifiers  
1..964  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="IMAGE:649441"  
/tissue\_type="retina"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_94"  
/note="Organ: eye; Vector: pCMV-Sport6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 3.3 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 21.5%; Score 775; DB 13; Length 964;  
Best Local Similarity 88.9%; Pred. No. 6.8e-185;  
Matches 857; Conservative 0; Mismatches 105; Indels 2; Gaps 2;  
Qy 1774 TGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCGGGGCTGGCAACTGATGAACAA 1833  
Db 1 TGTCTGTGGGACACACAGTGTCTCGATGTACCTCTGGGAGTTGGCACTGAAGACAA 60  
Qy 1834 GAAGAAAGTTAAATCAGATGTTTTCACAAAAGAACTCTTGACCAATGACAGATGTGAC 1893  
Db 61 GCAGAAAGTTAAATCAGAGTGTCTTCTAAAGAAACCTTGACCATGACAGATGTGAC 120  
Qy 1894 CAGCACACAGATTGTTACAGTGCACAGCCACCAACCAATGACTGCCACTGGTCAATGAC 1953  
Db 121 CAGCACACAGATTGTTACAGTGCACAGCCAAATACCAATGACTGCCACTGGTCAATGAT 180  
Qy 1954 CAATGTGTCCCGAGAAACCAAGCTGTCTCAGAAAGCCAGATCTCCATTTTAGTATGAG 2013  
Db 181 CACTGTGTCTCTGTGAACCCAGCTGCACAGAGCCAGATCTCCATTTGCCAGTATGAG 240  
Qy 2014 AATTGCCCCAAGGATAACCTTATGTACTGTAAAGAAAGACCACTGAGAGTGT 2073  
Db 241 AGTTGCCCCAAGGATAACCCCATGTACTGTCAATGAAGAAACCACTGAGAGTGT 300  
Qy 2074 GCGCTGAGACCAAGACTGCGAGTGGAGCCCGGAAATCAGAGTGCATTTGCCCTGCCGAA 2133  
Db 301 GCGCTGAGACCAAGACTGCGAGTGGAGCCCGGAAATCAGAGTGCATTTGCCCTGCCGAA 360  
Qy 2134 AATATCTGTGCAATTTGGTGGCAATTTGGTGGAAACCTCATGTTTGAATAATCTACTGCC 2193  
Db 361 AATATCTGTGCAATTTGGTGGCAATTTGGTGGAAACCTCATGTTTGAATAATCTACTGCC 420  
Qy 2194 AAGGAGATTATGCAATGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCTTCT 2253  
Db 421 AAGGAGATTATGCAATGCTAAATTTGTTCTGTAGGAACCAACATGCCCTTTGGCTTCT 480  
Qy 2254 CTTACACCCAGAAAGAGTAGAATTTGCTTAAAGCAGCTGCGAATTAATGCAAGTCACT 2313  
Db 481 CTCACATCCAGAAAGAGTAGAATTTGCTTAAAGCAGCTGCGAATTAATGCAAGTCACT 540  
Qy 2314 CAGAGCATGTCAGCTCACCTTAACCCCATGGCTGGCTTGGAGAGTCAATGCTGCC 2373  
Db 541 CAAGATATGTCACAGCTCACTCTGACTCCCATGGTGGTGGTCTTTCGGAAGATCAATGCTCT 600  
Qy 2374 TACTGGTCTCGGAGATATGTCCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCCTCT 2433  
Db 601 TACTGGTCTCGGAGATATGTCCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCCTCT 660  
Qy 2434 GAGCCAGTGTGCTGAGATTCTGTGGATTTTATCAGAACCCAGTACTC-GGGGACTGAA 2492  
Db 661 GAGCCAGTGTGCTGAGATTCTGTGGATTTTATCAGAACCCAGTACTC-GGGGACTGAA 720  
Qy 2493 GCGTCAACCTGCACTCAACCCACTCAATGTGTGTGTGTAAGAGGCTTCAACACCAAG 2552  
Db 721 GCGTCAACCTGCACTCAACCCACTCAATGTGTGTGTGTAAGAGGCTTCAACACCAAG 780

```
QY 2553 TGCTAGCAGTCCGACACCAATGTCCTTGAGGACAG-CATGTGGAGATTGACACAGG 2611
DB 781 TGCCAAGCAGTCCGACACCAATGTCCTTGCGGACANCTTTGTGGCCAGTGCATAGCA 840
QY 2612 CGAGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAGTGTGTGAGACTCCAATGCTATG 2671
DB 841 CCAGCTCGAGTGCATGTGGTGAGCAACATGAAGCAGTGTGTGAGACTCCAATGCTATG 900
QY 2672 TGCCCTCCCTCCCTTTGCGCAGTGTATGGAATGGATATGATAGAGACCTGCCCCCTG 2731
DB 901 TGCCCTCCCTCCCTTTGCGCAGTGTATGGAATGGATATGATAGAGACCTGCCAACTG 960
QY 2732 AAAA 2735
DB 961 AAAA 964

RESULT 6
BG678679
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

933 bp mRNA linear EST 01-MAY-2001
602624477F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',
mRNA sequence.
BG678679
602624477F1 GI:13910076
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@email.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10602 row: j column: 20
High quality sequence stop: 762.
Location/Qualifiers
1..933
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4749427"
/tissue_type="squamous cell carcinoma"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NCI_CGAP_Skn4"
/notes="Organ: skin; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.5kb. Library constructed by Life
Technologies. Note: this is a NCI_CGAP Library."

ORIGIN
Query Match 21.2%; Score 763.2; DB 12; Length 933;
Best Local Similarity 97.3%; Pred. No. 6.6e-182;
Matches 799; Conservative 0; Mismatches 18; Indels 4; Gaps 2;

QY 2775 ACCAGGCTGGTGGTGTACTGATCCAGCAATCTGCAAGGGAATGATAGAGGG 2834
DB 2 ACCAGGCTGGTGGTGTACTGATCCAGCAATCTGCAAGGGAATGATAGAGGG 61
QY 2835 TTCCCTATAGGACCACTGAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCCA 2894
DB 62 TTCCCTATAGGACCACTGAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCCA 121
QY 2895 GCCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCATACAATGCTTTTCATTC 2954
```

```
DB 122 GCCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCATACAATGCTTTTCATTC 181
QY 2955 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAA 3014
DB 182 CTGTCCAGCTTGCCAATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAA 241
QY 3015 GTGTGAGAACTGACCAACAGCAGCAGCACTGGGACCTGGATATCTGGCTTCTACGGTGA 3074
DB 242 GTGTGAGAACTGACCAACAGCAGCAGCACTGGGACCTGGATATCTGGCTTCTACGGTGA 301
QY 3075 TCCCACCAATGGAGGGAATGTTCAGCCATGCAAGTGGCAATGGGACGCTCTCTGTGCAA 3134
DB 302 TCCCACCAATGGAGGGAATGTTCAGCCATGCAAGTGGCAATGGGACGCTCTCTGTGCAA 361
QY 3135 CACCAACACGGGCAAGTGTCTTCACCAACCAAGGGGCTCAAGGGGACGAGTGCACGCT 3194
DB 362 CACCAACACGGGCAAGTGTCTTCACCAACCAAGGGGCTCAAGGGGACGAGTGCACGCT 421
QY 3195 ATGTGAGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATACTCT 3254
DB 422 ATGTGAGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATACTCT 481
QY 3255 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGC 3314
DB 482 TCTTATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGC 541
QY 3315 TATCAATTTTGTGCTACTCTCTGACGAACAAACAGGGATTGGACATGTTTCATCAATGC 3374
DB 542 TATCAATTTTGTGCTACTCTCTGACGAACAAACAGGGATTGGACATGTTTCATCAATGC 601
QY 3375 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGC 3434
DB 602 CTCCAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGC 661
QY 3435 TGGAGAGAGATGCTGCTGTTTCAAAAACCAACATTAAGAGGTACAAAGATGTTTCTC 3494
DB 662 TGGAGAGAGATGCTGCTGTTTCAAAAACCAACATTAAGAGGTACAAAGATGTTTCTC 721
QY 3495 TAATGAGAAGTTTGATTTTGCACACCCCAAAATATCACCTTTCTTTTGTATGTCAGTAA 3554
DB 722 TAATGAGAAGTTTGATTTTGCACACCCCAAAATATCACCTTTCTGTGTATGTCAGTAA 780
QY 3555 TTTCACCT---GGCCCATCAAAATTCAGGTGCAAACTGAA 3591
DB 781 TTTCACCTTGGGCCCATCGAAATTCAGGTGCAAATGGA 820

RESULT 7
CB520374
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

840 bp mRNA linear EST 09-JUL-2003
UI-M-GIO-cei-1-18-0-UI.r1 NIH_BMAP_GIO Mus musculus cDNA clone
IMAGE:6840067 5', mRNA sequence.
CB520374
CB520374.1 GI:29353729
EST.
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs@email.nih.gov
Tissue Procurement: Dr. Jim Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
```

```

(BMAP)
Seq primer: pyX-5.
Location/Qualifiers
  1..840
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL/6"
    /db_xref="taxon:10090"
    /clone="IMAGE:6840067"
    /tissue_type="whole brain"
    /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
    /lab_host="DH10B (T1 phage resistant)"
    /clone_lib="NIMH_BMAP_G10"
    /notes="Organ: Brain; Vector: pyX- Asc; Site 1: EcoR I;
    Site 2: Not I; The library was constructed according
    Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
    1996. Denatured RNA was size fractionated on a 1% agarose
    gel. First strand cDNA synthesis was primed with oligo-dT
    primer containing a Not I site. Double strand cDNA was
    size selected according to mRNA size fraction. Ligated
    with EcoR I adaptor, digested with NotI and then cloned
    directionally into pyX-Asc vector. The library tag
    sequence located between the Not I site and the polyA tail
    is AGCGAGACG. This library was created for the University
    Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
    Developing Mouse Nervous System', supported by National
    Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
    program coordinator.'
ORIGIN
Query Match          19.4%; Score 699.6; DB 14; Length 840;
Best Local Similarity 90.1%; Pred. No. 8.6e-166;
Matches 758; Conservative 0; Mismatches 82; Indels 1; Gaps 1;

QY      1929  CAATGACTGCCACTGGTGCATGACCATTTGTGTCCCGAGAACCAACGCTGCTCAGNAGG 1988
Db      1    CAATGACTGCCACTGGTGCATGACCATTTGTGTCCCGAGAACCAACGCTGCGACAGAAG 60

QY      1989  CCAGATCTCCATTTTAGGTATGAGAAATCCGCCAAGGATAACCCATGTACTACTGTAA 2048
Db      61    CCAGATCTCCATTTGCCAAGTATGAGATTGCCCCNAGGATAACCCCATGTACTACTGCAA 120

QY      2049  CAAGAAGACAGCTGCAGGAGCTGTGCCCTGGACCAAGAACTGCAGTGGAGGCCCGAA 2108
Db      121    TAAGAAAACAGCTGCAGGAGCTGTGCCCTAGACACAGAACTGCCAGTGGAGGCCCGAA 180

QY      2109  TCAGGAGTGCAATTTGCCCTGCCGAAATATCTGTGGCAATTTGGCTGGCATTTGTTGAAA 2168
Db      181    TCAAGAGTGCAATCGCCCTGCCGAAATATCTGTGGCAATTTGGCTGGCATTTGTTGAAA 240

QY      2169  CTCATGTTTGAAAATTACTACTGCCAGGAGAAATTATGACAAATGCTAAATTTGTTCTGTAG 2228
Db      241    CTCGTGTCTGAAAATCACTACTCTAAGGAGAAATTATGACAAATGCTAAATTTGTTCTGTAG 300

QY      2229  GAACCACAATGCCCTTTTGGCTTCTCTTACAAACCCAGAGAAGGTAGAAATTTGTCCTTAA 2288
Db      301    GAACCACAATTGCCTTTTGGCTTCTCTTACAAACCCAGAGAAGGTAGAAATTTGTCCTTAA 360

QY      2289  GCAGCTGCGAATAATGCAAGTCATCTCAGAGCATGTCCAAAGCTCACTTAAACCCCATGGGT 2348
Db      361    GCAGCTTCGATTAATGCAATCATCTCAAGTATGTCCAAAGCTCACTCTGACTCCATGGGT 420

QY      2349  CGGCTTTGGAAGATCAATGTGTCCTACTGTGTGCTGGGAAGATATGTCCTCCATTACAAA 2408
Db      421    TGTGTTCTTGGGAAGATCAATGTGTCCTACTGTGTGCTGGGAAGATATGTCCTCCATTACAAA 480

QY      2409  TAGTTTTACTACAGTGAATGCCGTCTGAGCCAGTGATGCTGGATTCCTGTGGAAATTTATC 2468
Db      481    TAGTTTTGCTCGATGGATGCAATCTGAGCCAGTGATGCTGGCTTCTGTGGATCTTGTC 540

QY      2469  AGAACCCAGTACTCGGGGACTGAAGGCTGCAACTGCAATCAACCCATCAATGGTAGTGT 2528
Db      541    AGAGCCTTAGTACTCGGGGATTAAGGCTGCAACTGCAATCAACCCATCAATGGTAGTGT 600

```

ORIGIN	method."	Unpublished (2002)
Query Match	18.9%; Score 681; DB 12; Length 681;	Contact: Kim YS
Best Local Similarity	100.0%; Pred. No. 48-161;	Genome Research Center
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Korea Research Institute of Bioscience & Biotechnology
		52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
		Tel: +82-42-860-4470
		Fax: +82-42-860-4409
		Email: yongsung@mail.kribb.re.kr
		Plate: 36 row: 5 column: 09
		High quality sequence stop: 683.
		Location/Qualifiers
		1. 683
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clones="SSSNU484-36-E09"
		/sex="M"
		/tissue_type="Stomach"
		/cell_line="Epithelial"
		/lab_host="Top10F"
		/clone_lib="SSSNU484"
		/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-tailed mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F' by electroporation method."
QY 2290	CAGCTGGGAATATGAGTCTCAGAGATGTCAGAGTCACTTAAACCCATGGGTC	
Db 1	CAGCTGGGAATATGAGTCTCAGAGATGTCAGAGTCACTTAAACCCATGGGTC	
QY 2350	GGCTTCGGGAATATGAGTCTCAGAGATGTCAGAGTCACTTAAACCCATGGGTC	
Db 61	GGCTTCGGGAATATGAGTCTCAGAGATGTCAGAGTCACTTAAACCCATGGGTC	
QY 2410	AGTTTACTACAGTGGATGCGCTGAGCCAGATGCTGGAGTCTGTGGAAATTTATCA	
Db 121	AGTTTACTACAGTGGATGCGCTGAGCCAGATGCTGGAGTCTGTGGAAATTTATCA	
QY 2470	GAACCCAGTACTCGGGGACTGAAGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGC	
Db 181	GAACCCAGTACTCGGGGACTGAAGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGC	
QY 2530	TGTGAAAGGCTGCAAAACACAGTCTGAAGCAGTCCGGACACCATGTGCTTGAAGACA	
Db 241	TGTGAAAGGCTGCAAAACACAGTCTGAAGCAGTCCGGACACCATGTGCTTGAAGACA	
QY 2590	GCATCTGGAGATTGACACAGCGGAGCTCTCAGTGTGCTGTCAGCAACATGAAGCAG	
Db 301	GCATCTGGAGATTGACACAGCGGAGCTCTCAGTGTGCTGTCAGCAACATGAAGCAG	
QY 2650	TGTGTGACTTCCAAATGCTATGTGGCTCTTCCCTTTGGCCAGTGTGAATGGTAT	
Db 361	TGTGTGACTTCCAAATGCTATGTGGCTCTTCCCTTTGGCCAGTGTGAATGGTAT	
QY 2710	ACGATGAGCACTGCCCCCTGAAAATGTTTCAGGCTACTTACTCTGTAGTCATTTGCTTG	
Db 421	ACGATGAGCACTGCCCCCTGAAAATGTTTCAGGCTACTTACTCTGTAGTCATTTGCTTG	
ORIGIN		
Query Match	18.9%; Score 680.4; DB 12; Length 683;	
Best Local Similarity	99.7%; Pred. No. 5.7e-161;	
Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
QY 2290	CAGCTGGGAATATGAGTCTCAGAGATGTCAGAGTCACTTAAACCCATGGGTC	2349
Db 1	CAGCTGGGAATATGAGTCTCAGAGATGTCAGAGTCACTTAAACCCATGGGTC	60
QY 2350	GGCTTCGGGAATATGAGTCTCAGAGATGTCAGAGTCACTTAAACCCATGGGTC	2409
Db 61	GGCTTCGGGAATATGAGTCTCAGAGATGTCAGAGTCACTTAAACCCATGGGTC	120
QY 2410	AGTTTACTACAGTGGATGCGCTGAGCCAGTGTGGAGTCTGTGGAAATTTATCA	2469
Db 121	AGTTTACTACAGTGGATGCGCTGAGCCAGTGTGGAGTCTGTGGAAATTTATCA	180
QY 2470	GAACCCAGTACTCGGGGACTGAAGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGC	2529
Db 181	GAACCCAGTACTCGGGGACTGAAGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGC	240
QY 2530	TGTGAAAGGCTGCAAAACACAGTCTGAAGCAGTCCGGACACCATGTGCTTGAAGACA	2589
Db 241	TGTGAAAGGCTGCAAAACACAGTCTGAAGCAGTCCGGACACCATGTGCTTGAAGACA	300
QY 2590	GCATCTGGAGATTGACACAGCGGAGCTCTCAGTGTGCTGTCAGCAACATGAAGCAG	2649
Db 301	GCATCTGGAGATTGACACAGCGGAGCTCTCAGTGTGCTGTCAGCAACATGAAGCAG	360
QY 2650	TGTGTGACTTCCAAATGCTATGTGGCTCTTCCCTTTGGCCAGTGTGAATGGTAT	2709
Db 361	TGTGTGACTTCCAAATGCTATGTGGCTCTTCCCTTTGGCCAGTGTGAATGGTAT	420
QY 2710	ACGATGAGCACTGCCCCCTGAAAATGTTTCAGGCTACTTACTCTGTAGTCATTTGCTTG	2769
Db 421	ACGATGAGCACTGCCCCCTGAAAATGTTTCAGGCTACTTACTCTGTAGTCATTTGCTTG	480
QY 2770	GAGCAACAGGCTGTGGCTGTACTGTATCCAGCAATATCGGCAAGGAAATGCATA	2829
Db 481	GAGCAACAGGCTGTGGCTGTACTGTATCCAGCAATATCGGCAAGGAAATGCATA	540
QY 2830	GAGGGTTCCTATAAGGACAGTGAAGATGCTTCCAGCCCTTACAGAAATTTCTAT	2889
Db 541	GAGGGTTCCTATAAGGACAGTGAAGATGCTTCCAGCCCTTACAGAAATTTCTAT	600
QY 2890	CCAGCCCTGCTCAATCCAGATGCTGTAGAGGACAGATCACTGGCTTTC	2949
Db 601	CCAGCCCTGCTCAATCCAGATGCTGTAGAGGACAGATCACTGGCTTTC	660
QY 2950	ATTCACGTGCTCCAGCTTGCCAA 2970	
Db 661	ATTCACGTGCTCCAGCTTGCCAA 681	
RESULT 9		
BM783788		
LOCUS	K-EST0061794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',	
DEFINITION	mRNA sequence.	
ACCESSION	BM783788.1 GI:19132020	
VERSION	EST.	
KEYWORDS	Homo sapiens (human)	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
REFERENCE	1 (bases 1 to 683)	
AUTHORS	Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,	
	Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and	
	Kim,Y.S.	
TITLE	21C Frontier Korean EST Project 2001	



```

QY 2770 GAGCAACAGGCTGTGGCTGTACTGATCCAGCAATACTGGCAAGGAAATGCATA 2829
Db 481 GAGCAACAGGCTGTGGCTGTACTGATCCAGCAATACTGGCAAGGAAATGCATA 540
QY 2830 GAGGTTTCTTATAAAGACACAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 2889
Db 541 GAGGTTTCTTATAAAGACACAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 600
QY 2890 CCACAGCCCTGTCTCAATTCAGCATGTCTAGAGACAGCAGATACAACTGGTCTTTC 2949
Db 601 CCACAGCCCTGTCTCAATTCAGCATGTCTAGAGACAGCAGATACAACTGGTCTTTC 660
QY 2950 ATTCACTGTCCAGCTTGCCAAATG 2972
Db 661 ATTCACTGTCCAGCTTGCCAAATG 683

RESULT 10
LOCUS BM783258
DEFINITION K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',
mRNA sequence.
ACCESSION BM783258
VERSION BM783258.1 GI:19131490
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 31 row: H column: 08
High quality sequence stop: 680.
Location/Qualifiers
1..680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-31-H08"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation
method."

FEATURES
source
1..680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-31-H08"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F' by electroporation
method."

ORIGIN
Query Match 18.9%; Score 680; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 7.2e-161;

```

```

Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2290 CAGCTGCGAAATAATGACAGTCAATCTCAGAGCAATGTCACAGCTCACTTAACCCCAATGGGTC 2349
Db 1 CAGCTGCGAAATAATGACAGTCAATCTCAGAGCAATGTCACAGCTCACTTAACCCCAATGGGTC 60
QY 2350 GGCCTTCGGAAGATCAATGTCTCTACTGTGTCGGGAAGATATGTCCCAATTTACAAAT 2409
Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGTGTCGGGAAGATATGTCCCAATTTACAAAT 120
QY 2410 AGTTTACTACAGTGAATGCCCTCTGAGCCCAAGTCAATGCTGGAATTTCTGGAATTTTATCA 2469
Db 121 AGTTTACTACAGTGAATGCCCTCTGAGCCCAAGTCAATGCTGGAATTTCTGGAATTTTATCA 180
QY 2470 GAACCCAGTACTCGGGGACTGAAGGCTGCACCTGCATCAACCCACATCAATGGTAGTGC 2529
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCACCTGCATCAACCCACATCAATGGTAGTGC 240
QY 2530 TGTGAAAGGCTGCAAAACCAACAGTCTAAGCAGTCCCGGACACCATGTCCTTGAGGACA 2589
Db 241 TGTGAAAGGCTGCAAAACCAACAGTCTAAGCAGTCCCGGACACCATGTCCTTGAGGACA 300
QY 2590 GCATGTGAGATTCGACACAGCGGAGCTCTGAGTGCATGTGTCGAGCAACATGAAGCAG 2649
Db 301 GCATGTGAGATTCGACACAGCGGAGCTCTGAGTGCATGTGTCGAGCAACATGAAGCAG 360
QY 2650 TGTGTGAGTCTCAATGCTTATGTGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
Db 361 TGTGTGAGTCTCAATGCTTATGTGCTCTTCCCTTTTGGCCAGTGTATGGAATGGTAT 420
QY 2710 ACGATGAGCACTGCGCCCTCGAAAAATGTTTCAGGCTACTGTACTGTAGTCAATGCTTG 2769
Db 421 ACGATGAGCACTGCGCCCTCGAAAAATGTTTCAGGCTACTGTACTGTAGTCAATGCTTG 480
QY 2770 GAGCAACAGGCTGTGGCTGTACTGATCCAGCAATACTGCCAAGGAAATGCATA 2829
Db 481 GAGCAACAGGCTGTGGCTGTACTGATCCAGCAATACTGCCAAGGAAATGCATA 540
QY 2830 GAGGTTTCTTATAAAGACACAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 2889
Db 541 GAGGTTTCTTATAAAGACACAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 600
QY 2890 CCACAGCCCTGTCTCAATTCAGCATGTCTAGAGACAGCAGATACAACTGGTCTTTC 2949
Db 601 CCACAGCCCTGTCTCAATTCAGCATGTCTAGAGACAGCAGATACAACTGGTCTTTC 660
QY 2950 ATTCACTGTCCAGCTTGCCA 2969
Db 661 ATTCACTGTCCAGCTTGCCA 680

RESULT 11
LOCUS BM783773
DEFINITION K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5',
mRNA sequence.
ACCESSION BM783773
VERSION BM783773.1 GI:19132005
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 680)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

```

```
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: D column: 02
High quality sequence stop: 680.
Location/Qualifiers
Source
1. .680
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-36-D02"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation
method."

FEATURES
Db 541 GAGGTTCTCTATAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 600
Qy 2890 CCACAGCCCTGCTCAATTCAGCATGTGTAGAGACAGACATACAACTGGTCTTTC 2949
Db 601 CCACAGCCCTGCTCAATTCAGCATGTGTAGAGACAGACATACAACTGGTCTTTC 660
Qy 2950 ATTCACTGTCAGCTTGCCA 2969
Db 661 ATTCACTGTCAGCTTGCCA 680

RESULT 12
BM783718
LOCUS K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5',
DEFINITION mRNA sequence.
ACCESSION BM783718
VERSION BM783718.1 GI:19131950
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 677)
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 34 row: F column: 01
High quality sequence stop: 677.
Location/Qualifiers
Source
1. .677
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-34-F01"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site_1: EcoRI;
Site_2: NotI; The poly (A)+ RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation
method."

ORIGIN
Query Match 18.9%; Score 677; DB 12; Length 680;
Best Local Similarity 100.0%; Pred. No. 7.2e-161;
Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2290 CAGCTGCGAATATGAGTATGCTCTACTGCTGCGGAGATATGTCCTTCAAAAT 2409
Db 1 CAGCTGCGAATATGAGTATGCTCTACTGCTGCGGAGATATGTCCTTCAAAAT 120
Qy 2350 GGCCTTCGGAGATCAATGTGCTCTACTGCTGCGGAGATATGTCCTTCAAAAT 2409
Db 61 GGCCTTCGGAGATCAATGTGCTCTACTGCTGCGGAGATATGTCCTTCAAAAT 120
Qy 2410 AGTTTACTACAGTGGATGCGCTGAGCCAGTGTGCTGGATCTGTGGAATTTATCA 2469
Db 121 AGTTTACTACAGTGGATGCGCTGAGCCAGTGTGCTGGATCTGTGGAATTTATCA 180
Qy 2470 GAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCACTCAATGCTAGTGTG 2529
Db 181 GAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCACTCAATGCTAGTGTG 240
Qy 2530 TGTGAAGGCTGCAACCACTGCTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2589
Db 241 TGTGAAGGCTGCAACCACTGCTGAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 300
Qy 2590 GCATGTGAGATTCACACGCGCAGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2649
Db 301 GCATGTGAGATTCACACGCGCAGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 360
Qy 2650 TGTGTGAGTCCAAATGCTGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTAT 2709
Db 361 TGTGTGAGTCCAAATGCTGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTAT 420
Qy 2710 ACGATGAGCAGCTGCCCCCTGAAATGTTGAGCTACTGAGCTGTAGTCTGCTGCTG 2769
Db 421 ACGATGAGCAGCTGCCCCCTGAAATGTTGAGCTACTGAGCTGTAGTCTGCTGCTG 480
Qy 2770 GAGCAACGAGCTGTGGCTGTGTACTGTATCCAGCAATATCGGCAAGGGAATGCATA 2829
Db 481 GAGCAACGAGCTGTGGCTGTGTACTGTATCCAGCAATATCGGCAAGGGAATGCATA 540
Qy 2830 GAGGGTTCCTATAAGGACCAAGTGAAGATGCTTTCGCAAGCCCTCAGGAAATTTCTAT 2889
```

```

QY 2350 GGCCTTCGGAAGATCAATGTGTCTTACTGTGTGGGAAGATATGTCCTCCCATTTACAAAT 2409
Db 61 GGCCTTCGGAAGATCAATGTGTCTTACTGTGTGGGAAGATATGTCCTCCCATTTACAAAT 120
QY 2410 AGTTTACTACAGTGGATGCGCTGTAGCCAGTATGCTGGATCTGTGGAAATTTTATCA 2469
Db 121 AGTTTACTACAGTGGATGCGCTGTAGCCAGTATGCTGGATCTGTGGAAATTTTATCA 180
QY 2470 GAACCCAGTACTCGGGGACTGAAGGCTCAACCTGCAATCAACCACTCAATGGTAGTGTC 2529
Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAATCAACCACTCAATGGTAGTGTC 240
QY 2530 TGTGAAGGCTGCAAAACCAACAGTGAAGTATGCGGACACCATGTGCTTGGAGACA 2589
Db 241 TGTGAAGGCTGCAAAACCAACAGTGAAGTATGCGGACACCATGTGCTTGGAGACA 300
QY 2590 GCATGTGGAGATTCACACAGCGGAGCTCTGAGTGCATGTTGGTGCAGCAACATGAAGCAG 2649
Db 301 GCATGTGGAGATTCACACAGCGGAGCTCTGAGTGCATGTTGGTGCAGCAACATGAAGCAG 360
QY 2650 TGTGTGAGCTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAAATGCTAT 2709
Db 361 TGTGTGAGCTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAAATGCTAT 420
QY 2710 ACGATGAGCACTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGTAGTCAATGCTTG 2769
Db 421 ACGATGAGCACTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGTAGTCAATGCTTG 480
QY 2770 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGAAATACTTGGCAAGGAAATGCATA 2829
Db 481 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATACTTGGCAAGGAAATGCATA 540
QY 2830 GAGGTTTCTATAAAGGACCAAGTCAAGATGCTTCGCAAGCCCCCTACAGAAATTTCTAT 2889
Db 541 GAGGTTTCTATAAAGGACCAAGTCAAGATGCTTCGCAAGCCCCCTACAGAAATTTCTAT 600
QY 2890 CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949
Db 601 CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660
QY 2950 ATTCACTGTCCAGTTG 2966
Db 661 ATTCACTGTCCAGTTG 677

RESULT 13
BM783821
LOCUS K-EST0061830 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-H09 5',
DEFINITION mRNA sequence.
ACCESSION BM783821
VERSION 1 GI:19132053
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 677)
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,
Oh,K.J., Cheong,J.B., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and
Kim,Y.S.
TITLE 21C Frontier Korean EST Project 2001
JOURNAL Unpublished (2002)
COMMENT Contact: Kim YS
Genome Research Center
Korea Research Institute of Bioscience & Biotechnology
52 Eeun-dong Yuseong-gu, Daejeon 305-333, South Korea
Tel: +82-42-860-4470
Fax: +82-42-860-4409
Email: yongsung@mail.kribb.re.kr
Plate: 36 row: H column: 09
High quality sequence stop: 677.

```

## FEATURES

Location/Qualifiers  
 1..677  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clones="S5SNU484-36-H09"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /cell\_line="SNU-484"  
 /lab\_host="Top10P"  
 /clone\_lib="S5SNU484"  
 /note="Organ: Stomach; Vector: pTZ19BP1; Site 1: EcoRI;  
 Site 2: NotI; The poly (A)+ RNA was decapped with tobacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dt-selected  
 mRNA by priming with dt-tailed vector. The dt-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10P by electroporation  
 method."

## ORIGIN

Query Match 18.7%; Score 674.4; DB 12; Length 677;  
 Best Local Similarity 99.7%; Pred. No. 1.9e-159;  
 Matches 675; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 2290 CAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAGTCACTTAACCCCATGGGTC 2349  
 Db 1 CAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAGTCACTTAACCCCATGGGTC 60  
 QY 2350 GGCCTTCGGAAGATCAATGTGTCTTACTGTGTGGGAAGATATGTCCTCCCATTTACAAAT 2409  
 Db 61 GGCCTTCGGAAGATCAATGTGTCTTACTGTGTGGGAAGATATGTCCTCCCATTTACAAAT 120  
 QY 2410 AGTTTACTACAGTGGATGCGCTGTAGCCAGTATGCTGGATCTGTGGAAATTTTATCA 2469  
 Db 121 AGTTTACTACAGTGGATGCGCTGTAGCCAGTATGCTGGATCTGTGGAAATTTTATCA 180  
 QY 2470 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCACTCAATGGTAGTGTC 2529  
 Db 181 GAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCACTCAATGGTAGTGTC 240  
 QY 2530 TGTGAAGGCTGCAAAACCAACAGTGAAGTATGCGGACACCATGTGCTTGGAGACA 2589  
 Db 241 TGTGAAGGCTGCAAAACCAACAGTGAAGTATGCGGACACCATGTGCTTGGAGACA 300  
 QY 2590 GCATGTGGAGATTCACACAGCGGAGCTCTGAGTGCATGTTGGTGCAGCAACATGAAGCAG 2649  
 Db 301 GCATGTGGAGATTCACACAGCGGAGCTCTGAGTGCATGTTGGTGCAGCAACATGAAGCAG 360  
 QY 2650 TGTGTGAGCTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAAATGCTAT 2709  
 Db 361 TGTGTGAGCTCCATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAAATGCTAT 420  
 QY 2710 ACGATGAGCACTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGTAGTCAATGCTTG 2769  
 Db 421 ACGATGAGCACTGCCCCCTGAAAATTTGTCAGGCTACTGTACTGTAGTCAATGCTTG 480  
 QY 2770 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGAAATACTTGGCAAGGAAATGCATA 2829  
 Db 481 GAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATACTTGGCAAGGAAATGCATA 540  
 QY 2830 GAGGTTTCTATAAAGGACCAAGTCAAGATGCTTCGCAAGCCCCCTACAGAAATTTCTAT 2889  
 Db 541 GAGGTTTCTATAAAGGACCAAGTCAAGATGCTTCGCAAGCCCCCTACAGAAATTTCTAT 600  
 QY 2890 CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTC 2949

Db 601 CCACAGCCCTGCTCAATCCAGCATGTGCTAGAGGACAGACAGATACAACCTGGTCTTTC 660  
 QY 2950 ATTCACTGTCACGCTTG 2966  
 Db 661 ATTCACTGTCACGCTTG 677

RESULT 14  
 BM783685 673 bp mRNA linear EST 05-MAR-2002  
 LOCUS X-EST0001659 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-B06 5',  
 DEFINITION mRNA sequence.

ACCESSION BM783685  
 VERSION 1 GI:19131917  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 673)  
 AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001  
 JOURNAL Unpublished (2002)  
 COMMENT Contact: Kim YS  
 Korea Research Center  
 52 Eeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
 Tel: +82-42-860-4470  
 Fax: +82-42-860-4409  
 Email: yongseung@mail.kribb.re.kr  
 Plate: 34 row: B column: 06  
 High quality sequence stop: 673.

FEATURES  
 1..673  
 Location/Qualifiers

source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="S5SNU484-34-B06"  
 /sex="M"  
 /tissue\_type="Stomach"  
 /cell\_type="Epithelial"  
 /cell\_line="SNU-484"  
 /lab\_host="Top10F"  
 /clone\_lib="S5SNU484"  
 /note="Organ: Stomach; Vector: pTZ19RP1; Site 1: EcoRI;  
 Site 2: NotI; The poly (A)+ RNA was decapped with tabacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dT-selected  
 mRNA by priming with dT-tailed vector. The dT-tailed  
 vector was adjusted to have about 60nt. The cDNA vector  
 was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a cDNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transfection of  
 competent cells E. coli Top10F' by electroporation  
 method."

ORIGIN  
 Query Match 18.7%; Score 673; DB 12; Length 673;  
 Best Local Similarity 100.0%; Pred. No. 4,3e-159;  
 Matches 673; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2290 CAGCTGGATATGCAAGTCAATCTCAGAGCATGTCAGAGTCACCTTAACCCCATGGGTC 2349  
 Db 1 CAGCTGGATATGCAAGTCAATCTCAGAGCATGTCAGAGTCACCTTAACCCCATGGGTC 60

QY 2350 GGCCTTCGGAGATCAATGTGCTCTACTGGTCTGGAGATATGCCCATTTACAAT 2409  
 Db 61 GGCCTTCGGAGATCAATGTGCTCTACTGGTCTGGAGATATGCCCATTTACAAT 120

QY 2410 AGTTTACTACAGTGGATGCGCTCTGAGCCAGTCATGCTGGATTTCTTGGAAATTTATCA 2469  
 Db 121 AGTTTACTACAGTGGATGCGCTCTGAGCCAGTCATGCTGGATTTCTTGGAAATTTATCA 180

QY 2470 GAACCCAGTACTCGGGAGCTGAAGCTGGAACCTGCAATCAACCCACTCAATGGTAGTGC 2529  
 Db 181 GAACCCAGTACTCGGGAGCTGAAGCTGGAACCTGCAATCAACCCACTCAATGGTAGTGC 240

QY 2530 TGTGAAGGCTGCAAAACCAACAGTCTAAGCAGTGGCGGACACCATGTGCTTGGAGACA 2589  
 Db 241 TGTGAAGGCTGCAAAACCAACAGTCTAAGCAGTGGCGGACACCATGTGCTTGGAGACA 300

QY 2590 GCATGTGGAGATTGCACACGCGGAGCTCTGAGTGGCATGTGGTGACAGCAACATGAAGCAG 2649  
 Db 301 GCATGTGGAGATTGCACACGCGGAGCTCTGAGTGGCATGTGGTGACAGCAACATGAAGCAG 360

QY 2650 TGTGTGAGCTCCAATGCTATGTGCTCTCTCCCTTTTGGCCAGTGTATGGAATGTAT 2709  
 Db 361 TGTGTGAGCTCCAATGCTATGTGCTCTCTCCCTTTTGGCCAGTGTATGGAATGTAT 420

QY 2710 ACGATGAGCAGCTCCGCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCATGCTTG 2769  
 Db 421 ACGATGAGCAGCTCCGCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCATGCTTG 480

QY 2770 GAGCAACGAGCTGTGGTGTGTACTGATCCAGCAATACTGCAAGGGAATGCATA 2829  
 Db 481 GAGCAACGAGCTGTGGTGTGTACTGATCCAGCAATACTGCAAGGGAATGCATA 540

QY 2830 GAGGTTCTCTATAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 2889  
 Db 541 GAGGTTCTCTATAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT 600

QY 2890 CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGCAGCAGATCAACTGCTTTTC 2949  
 Db 601 CCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGCAGCAGATCAACTGCTTTTC 660

QY 2950 ATTCACTGTCACG 2962  
 Db 661 ATTCACTGTCACG 673

RESULT 15  
 BM783685 962 bp mRNA linear EST 15-MAY-2003  
 LOCUS BX431593  
 DEFINITION CS0DF012YU19 5-PRIME, mRNA sequence.

ACCESSION BX431593  
 VERSION BX431593  
 KEYWORDS EST.

SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 962)  
 AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 10212.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0BAG0172B02\_CS01546\_1&cluster=10212.f.  
 Contact : Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0BAG0172B02\_CS01546\_1.

FEATURES  
 1..962  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"

/db\_xref="taxon:9606"  
/clone="CS0DF012YL19"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end  
enriched, double-strand cDNA was digested with Not I and  
cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
vector. Library was not normalized."

## ORIGIN

Query Match 18.6%; Score 669.8; DB 13; Length 962;  
Best Local Similarity 99.4%; Pred. No. 3.4e-158;  
Matches 697; Conservative 0; Mismatches 8; Indels 3; Gaps 2;

Qy 2891 CACAGCCCTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATA-CAACTGGTCTTTTC 2949  
Db 7 CACAGCCCTGCT--AATTCAGATGTGTCTAGAGACAGCAGATACCAACTGGTCTTTTC 64

Qy 2950 ATTCACTGTCCAGTTGCCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGT 3009  
Db 65 ATTCACTGTCCAGTTGCCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGT 124

Qy 3010 GAGAAGTGTGAGAACCTTGACACAGGCAAGCACTGCGAGACCTGCGATATCTGGCTTCTTAC 3069  
Db 125 GAGAAGTGTGAGAACCTTGACACAGGCAAGCACTGCGAGACCTGCGATATCTGGCTTCTTAC 184

Qy 3070 GGTGATCCCAACCAATGAGGGAAATGTGAGCCATGCAAGTGCAATGGGCGACGGCTCTCTG 3129  
Db 185 GGTGATCCCAACCAATGAGGGAAATGTGAGCCATGCAAGTGCAATGGGCGACGGCTCTCTG 244

Qy 3130 TGCAACACCAACACGGGCAAGTGTCTCTGACCAACCAAGGGCGTCAAGGGGGACGAGTGC 3189  
Db 245 TGCAACACCAACACGGGCAAGTGTCTCTGACCAACCAAGGGCGTCAAGGGGGACGAGTGC 304

Qy 3190 CAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTAT 3249  
Db 305 CAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTAT 364

Qy 3250 ACTCTTCTTATTGACTATGATTTCACTTTAGTCTATCCAGGAGAGATCGCTATTATAC 3309  
Db 365 ACTCTTCTTATTGACTATGATTTCACTTTAGTCTATCCAGGAGAGATCGCTATTATAC 424

Qy 3310 ACAGCTATCAATTTTGTGGTACTCTCTGAGAAACAAAACAGGGATTTGGACATGTTTCATC 3369  
Db 425 ACAGCTATCAATTTTGTGGTACTCTCTGAGAAACAAAACAGGGATTTGGACATGTTTCATC 484

Qy 3370 AATGCTTCCAGAAATTTCAACATCACTCGGCTGCCAGTTTCTCAGCTGGAACC 3429  
Db 485 AATGCTTCCAGAAATTTCAACATCACTCGGCTGCCAGTTTCTCAGCTGGAACC 544

Qy 3430 CAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 3489  
Db 545 CAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGT 604

Qy 3490 TTCTTAATCAGAAAGTTGATTTTCGCAACACCCAAAATATCACTTTCTTTGTTTATGTC 3549  
Db 605 TTCTTAATCAGAAAGTTGATTTTCGCAACACCCAAAATATCACTTTCTTTGTTTATGTC 664

Qy 3550 AGTAATTTCACTGGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597  
Db 665 AGTAATTTCACTGGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 712

Search completed: March 5, 2004, 12:07:51  
Job time : 5707.24 secs

Result No.	Query			ID	Description
	Score	Match	Length		
1	3597	100.0	3597	3	Aas252274 Human sol
2	3597	100.0	3597	3	Aas72657 DNA encod
3	3597	100.0	3597	9	Ad470997 Human att
4	3556.4	98.9	4068	3	Aas52275 Human mem
5	3494.4	97.1	4072	3	Aas252275 Human mem
6	3487	96.9	3819	3	Aas252277 Human sol
7	3476.8	96.7	3850	3	Aas252277 Human sol
8	3469.4	96.5	4290	3	Aas52276 Human mem
9	3469.4	96.5	8561	5	Aas72659 DNA encod
10	2909	80.9	8739	9	Adb53249 Primary r
11	2843.6	79.1	8827	3	Aas291816 Wild type
12	2790.4	77.6	3490	5	Aas72860 DNA encod
13	2028.8	56.4	2625	3	Aas291322 Human mah
14	1428.8	39.7	2419	3	Aas291917 Murine mah
15	1341.6	37.3	6373	3	Aas291919 Human mah
16	1212.2	37.7	4134	7	Ab258710 Murine hom
17	1203.2	33.5	4140	7	Ab258709 Human hom
18	1129.4	31.4	3609	6	Ab563434 Human cdn
19	873.2	24.3	3165	5	Aas78781 DNA encod
20	873.2	24.3	3313	5	Aas76294 DNA encod
21	817.8	22.7	1012	2	Aas84568 Biorhythm
22	574	16.0	6517	4	Aah98748 Human BSR
23	536	14.9	6532	4	Aas86766 DNA encod



CC The present sequence is a cDNA encoding soluble attractin-1 obtained from  
CC activated T cell and human foetal liver libraries. This sequence can be  
CC used to enhance immune response in immunosuppressed patients such as  
CC those undergoing chemo- and radio-therapy treatment for cancer or those  
CC suffering from common variable immunodeficiency syndrome. The proteins  
CC may also be used to screen modulators (agonists and antagonists) of  
CC immune responses which may also be used to regulate immune reactions.  
CC Attractin antibodies can be used to inhibit immune response in transplant  
CC recipients or patients afflicted with autoimmune disease  
XX

Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match	100.0%;	Score 3597;	DB 3;	Length 3597;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3597;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

1 ATGTTGGCCGAGCGGCGCAACTGAGGCAAGCTGAGGAGGAGGACGGCGGCGAGCGCA 60  
1 ATGTTGGCCGAGCGGCGCAACTGAGGCAAGCTGAGGAGGAGGAGCGGCGGCGAGCGCA 60  
61 GCGCTCGCGGCGAGGAGCGGCGCGCACTGTGTCAAACGCGCGTGCCTGCAACCCCTGGC 120  
61 GCGCTCGCGGCGAGGAGCGGCGCGCACTGTGTCAAACGCGCGTGCCTGCAACCCCTGGC 120  
121 ACCGCGCAGTGCCTGCTCCCGCGCTGGGTGGGCGAGCAATGCCAGCACTGCCGGGGGC 180  
121 ACCGCGCAGTGCCTGCTCCCGCGCTGGGTGGGCGAGCAATGCCAGCACTGCCGGGGGC 180  
181 CGCTTCAGACTAATGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATATAAATAC 240  
181 CGCTTCAGACTAATGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATATAAATAC 240  
241 AAAACGAAGTCAGCTGGCTCATTAAGGACAGCAATAGATAATGAGACTTCGTTTC 300  
241 AAAACGAAGTCAGCTGGCTCATTAAGGACAGCAATAGATAATGAGACTTCGTTTC 300  
301 AATCATTTTGTCTACAGAGTGTAGTTGGACCAATTTATGTTTATGATGGGACTCAATT 360  
301 AATCATTTTGTCTACAGAGTGTAGTTGGACCAATTTATGTTTATGATGGGACTCAATT 360  
361 TATGCAACGCTAGTGTGCTCAATTTAGTGGCTCAATTTGCTCCTGAGAGATGGCAATGAG 420  
361 TATGCAACGCTAGTGTGCTCAATTTAGTGGCTCAATTTGCTCCTGAGAGATGGCAATGAG 420  
421 ACTGTCCCTGAGTGTGTTGCCACATCAGGTTATGCTTGTGCTGATTTTATGATGATGCT 480  
421 ACTGTCCCTGAGTGTGTTGCCACATCAGGTTATGCTTGTGCTGATTTTATGATGATGCT 480  
481 GCTTATTAATTTGACTGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 540  
481 GCTTATTAATTTGACTGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAAT 540  
541 TCAGGCGGAGGAGTGTAGATCAGTAAATAGCAGGATCTCTGAATGTGATGCTCT 600  
541 TCAGGCGGAGGAGTGTAGATCAGTAAATAGCAGGATCTCTGAATGTGATGCTCT 600  
601 GAAACCTGGAAGAGTGAAGCATGTGACATTCCTCAGTGTACAGCAACTGTGTTTCTCT 660  
601 GAAACCTGGAAGAGTGAAGCATGTGACATTCCTCAGTGTACAGCAACTGTGTTTCTCT 660  
661 CATCGAGCATCTGCAATTAAGTGTGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720  
661 CATCGAGCATCTGCAATTAAGTGTGTCAGAGGATGCTCCTGCTTCTCAGACTGGCAG 720  
721 GGTCTCGATGTTTCAGTTCCTGTACCACTAAGCACTCAATTTGGACTCGAGGAAATAT 780  
721 GGTCTCGATGTTTCAGTTCCTGTACCACTAAGCACTCAATTTGGACTCGAGGAAATAT 780  
781 TCTAACTTAAAGTCTCCCGAGCATCTCATTAAGTGTGCTCAATGGAACATTTATGG 840  
781 TCTAACTTAAAGTCTCCCGAGCATCTCATTAAGTGTGCTCAATGGAACATTTATGG 840  
841 GTTGTGGAGATATATGTTCAACCACTACAGATTAATTAATTAATTAATTAATTAATTAAT 900

841	GTGTTGGAGATATATGTTCAACCACTACAGATTATAACATGGTTCTAGCGTATGACCTT	900
901	GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTCTGAACAATGTGGTTGTTAGATATGGT	960
901	GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTCTGAACAATGTGGTTGTTAGATATGGT	960
961	CATTCTTTGGCATTATACAGGATAAAATTTACATGTATGAGGAGGAAAAATGATTCACT	1020
961	CATTCTTTGGCATTATACAGGATAAAATTTACATGTATGAGGAGGAAAAATGATTCACT	1020
1021	GGGATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGATCATGGGTGTTCTTG	1080
1021	GGGATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGATCATGGGTGTTCTTG	1080
1081	ACCCCTAAGGCAAAAGGAGCAGTATGCAAGTGGTTGGGCACTCTGCACACATTTGTACACTG	1140
1081	ACCCCTAAGGCAAAAGGAGCAGTATGCAAGTGGTTGGGCACTCTGCACACATTTGTACACTG	1140
1141	AAGATGGCCGAGTGGTCTGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATATA	1200
1141	AAGATGGCCGAGTGGTCTGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATATA	1200
1201	AGCAATGTGACAGGAATATGATTTGGATAAAGAACACATGGAGTATATTACACACCCAGGGT	1260
1201	AGCAATGTGACAGGAATATGATTTGGATAAAGAACACATGGAGTATATTACACACCCAGGGT	1260
1261	GCCCTTGTGAAGGGGTTACGGCCATAGCAGTGTTCACGACCATAGGACAGGCGCCCTA	1320
1261	GCCCTTGTGAAGGGGTTACGGCCATAGCAGTGTTCACGACCATAGGACAGGCGCCCTA	1320
1321	TACGTTCAATGTGCTCAAGGCTTTCAAGTGCCTAAAGTACCGGCTTGAGATGATCTC	1380
1321	TACGTTCAATGTGCTCAAGGCTTTCAAGTGCCTAAAGTACCGGCTTGAGATGATCTC	1380
1381	TACGATATGATGTGGATACCCAGATGTGGACCAATCTTAAAGGACACCCGATTTTCCTG	1440
1381	TACGATATGATGTGGATACCCAGATGTGGACCAATCTTAAAGGACACCCGATTTTCCTG	1440
1441	TACTTGACACACAGCTGTGATAGTGGAGCAACATGCTGGTGTGGGGGAAACACACAC	1500
1441	TACTTGACACACAGCTGTGATAGTGGAGCAACATGCTGGTGTGGGGGAAACACACAC	1500
1501	AATGACACATCTATGAGCCATGGCCCAAAATGCTTCTTTCAGATTTTCATGGCTATGAC	1560
1501	AATGACACATCTATGAGCCATGGCCCAAAATGCTTCTTTCAGATTTTCATGGCTATGAC	1560
1561	ATTGCTGTGACCCGCTGGTCAAGTGTCCAGACCTGATGATGTTTGGTGGTTCAATAGT	1620
1561	ATTGCTGTGACCCGCTGGTCAAGTGTCCAGACCTGATGATGTTTGGTGGTTCAATAGT	1620
1621	TTGGCCATTTCAGCAGCTTTACACACACAGCACCATTGATGTTTGGTGGTTCAATAGT	1680
1621	TTGGCCATTTCAGCAGCTTTACACACACAGCACCATTGATGTTTGGTGGTTCAATAGT	1680
1681	CTCTCTCTCAGCAGATCTCTGTTTACCTCGAAACAGTGTGATGCGCATCGAGGTGAA	1740
1681	CTCTCTCTCAGCAGATCTCTGTTTACCTCGAAACAGTGTGATGCGCATCGAGGTGAA	1740
1741	GCCGCTTCTTTAGCAGGAGGACCTGCTATTCGTTGTGTTGGAAACACAGGGTCTCTCAG	1800
1741	GCCGCTTCTTTAGCAGGAGGACCTGCTATTCGTTGTGTTGGAAACACAGGGTCTCTCAG	1800
1801	TGTATCTCGTGGCGCTGGCAACTGATGAAACAGGAAAGTTAAATCAGATGTTTTT	1860
1801	TGTATCTCGTGGCGCTGGCAACTGATGAAACAGGAAAGTTAAATCAGATGTTTTT	1860
1861	TCCAAAAGAACTCTTTGACCATGACAGATGTACACAGACACAGATTTGTACAGCTGACA	1920
1861	TCCAAAAGAACTCTTTGACCATGACAGATGTACACAGACACAGATTTGTACAGCTGACA	1920
1921	GCCAAACCAATGATGCTGCCACTGGTGAATGACCAATTTGTGTCGCCAGAACACAGCTGC	1980

Db 1921 GCCAACACCAATGACTGGCCACTGGTCAATGACCAATTTGTGTCCCGAGAACACACAGCTGC 1980  
QY 1981 TCAGAGGCCAGATCTCCATTTTAGTATGAGAAATGGCCCAAGATTAACCCCTATGTAC 2040  
Db 1981 TCAGAGGCCAGATCTCCATTTTAGTATGAGAAATGGCCCAAGATTAACCCCTATGTAC 2040  
QY 2041 TACTGTAAACAGAGACCACTGACAGAGCTGGCCCTGGACACAGAACTGCCAGTGGGAG 2100  
Db 2041 TACTGTAAACAGAGACCACTGACAGAGCTGGCCCTGGACACAGAACTGCCAGTGGGAG 2100  
QY 2101 CCCCGAATCAGAGTGCATTTGCCCTGGCCGGAATAATCTGTGGCAATGGCTGGCAATTG 2160  
Db 2101 CCCCGAATCAGAGTGCATTTGCCCTGGCCGGAATAATCTGTGGCAATGGCTGGCAATTG 2160  
QY 2161 GTTGGAACTCATGTTTGAATAATCTACTGCGCAAGAGAAATTTATGACAAATGCTAAATTG 2220  
Db 2161 GTTGGAACTCATGTTTGAATAATCTACTGCGCAAGAGAAATTTATGACAAATGCTAAATTG 2220  
QY 2221 TTCTGTAGGAACACCAATGCCCTTTTGGCTTCTTTTACACCCAGAAAGGTAGAAATTT 2280  
Db 2221 TTCTGTAGGAACACCAATGCCCTTTTGGCTTCTTTTACACCCAGAAAGGTAGAAATTT 2280  
QY 2281 GTCCCTTAAGCAGCTCGGAATATGCAATGATCTCAGAGCATGTCGAGCTCACTTAACC 2340  
Db 2281 GTCCCTTAAGCAGCTCGGAATATGCAATGATCTCAGAGCATGTCGAGCTCACTTAACC 2340  
QY 2341 CCATGGGTGGCCCTTCGGAGATCAATGTGTCTCTACTGGTGTGGGAAGATATGTCCTCA 2400  
Db 2341 CCATGGGTGGCCCTTCGGAGATCAATGTGTCTCTACTGGTGTGGGAAGATATGTCCTCA 2400  
QY 2401 TTTTCAAAATAGTTTACTACAGTGGATGCCCTGTGAGCCAGCATGATCTGTGGA 2460  
Db 2401 TTTTCAAAATAGTTTACTACAGTGGATGCCCTGTGAGCCAGCATGATCTGTGGA 2460  
QY 2461 ATTTTATCAAAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCAATCAACCCACTCAAT 2520  
Db 2461 ATTTTATCAAAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCAATCAACCCACTCAAT 2520  
QY 2521 GGTAGTGTCTGTAAGGCTGCAAAACCAAGTGTAAAGCTGCGGACACCAATGTGCC 2580  
Db 2521 GGTAGTGTCTGTAAGGCTGCAAAACCAAGTGTAAAGCTGCGGACACCAATGTGCC 2580  
QY 2581 TTGAGGACAGCATGTGGAGATTCACACGGGCGAGCTCTGAGTGCATGTGGTGCAGCAAC 2640  
Db 2581 TTGAGGACAGCATGTGGAGATTCACACGGGCGAGCTCTGAGTGCATGTGGTGCAGCAAC 2640  
QY 2641 ATGAAGCAGTGTGGACTCCAAATGCTATGTGGCCCTCTTCCCTTTTGGCCAGTGTATG 2700  
Db 2641 ATGAAGCAGTGTGGACTCCAAATGCTATGTGGCCCTCTTCCCTTTTGGCCAGTGTATG 2700  
QY 2701 GAATGTATACGATGAGCAGCTGCCCTGGAATAATGTCAGGCTACTGTACCTGTAGT 2760  
Db 2701 GAATGTATACGATGAGCAGCTGCCCTGGAATAATGTCAGGCTACTGTACCTGTAGT 2760  
QY 2761 CAATGCTTGAGCAACACAGGCTGTGGTGTGCTGATCTCCAGCAATATGCGCAAGGG 2820  
Db 2761 CAATGCTTGAGCAACACAGGCTGTGGTGTGCTGATCTCCAGCAATATGCGCAAGGG 2820  
QY 2821 AATGCAATAGAGGTTCTTATAAGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGA 2880  
Db 2821 AATGCAATAGAGGTTCTTATAAGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGA 2880  
QY 2881 AATTTCTATCCAGCCCTGCTCAATTCAGCATGTCTAGAGGACAGCATACAC 2940  
Db 2881 AATTTCTATCCAGCCCTGCTCAATTCAGCATGTCTAGAGGACAGCATACAC 2940  
QY 2941 TGGTCTTTCAATCACTGTCCAGTTCGCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000  
Db 2941 TGGTCTTTCAATCACTGTCCAGTTCGCAATGCAACGGCCACAGTAAATGCATCAATCAG 3000  
QY 3001 AGCATCTGTAGAGTGTGAGACCTGACACAGGCAAGCACTGCGAGACCTGCATATCT 3060  
Db 3001 AGCATCTGTAGAGTGTGAGACCTGACACAGGCAAGCACTGCGAGACCTGCATATCT 3060

QY 3061 GGCTTCTACGCTGATCCCAATGGAGGAAATGTGAGCCATGCAATGCAATGGGCAC 3120  
Db 3061 GGCTTCTACGCTGATCCCAATGGAGGAAATGTGAGCCATGCAATGCAATGGGCAC 3120  
QY 3121 GCGTCTCTGTGCAACACCAACACCGGGAAGTGTCTTCTGACCAACCAAGGGCGTCAAGGGG 3180  
Db 3121 GCGTCTCTGTGCAACACCAACACCGGGAAGTGTCTTCTGACCAACCAAGGGCGTCAAGGGG 3180  
QY 3181 GACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAAACCCCTCTCAGAGGAACA 3240  
Db 3181 GACGAGTCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAAACCCCTCTCAGAGGAACA 3240  
QY 3241 TGTATTATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGAT 3300  
Db 3241 TGTATTATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGAT 3300  
QY 3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCTGACGAAACAAAACAGGGATTTGGAC 3360  
Db 3301 CGCTATTACACAGCTATCAATTTTGTGGCTACTCTGACGAAACAAAACAGGGATTTGGAC 3360  
QY 3361 ATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGCTGCCAGTTTCTCA 3420  
Db 3361 ATGTTTCATCAATGCTCCAGAAATTTCAACCTCAACATCACCTGGCTGCCAGTTTCTCA 3420  
QY 3421 GCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTAC 3480  
Db 3421 GCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTAC 3480  
QY 3481 AAGATAGTTTCTCTAATGAGAAGTTTGAATTTTGGCAACCAACCAATATCACTTTCTTT 3540  
Db 3481 AAGATAGTTTCTCTAATGAGAAGTTTGAATTTTGGCAACCAACCAATATCACTTTCTTT 3540  
QY 3541 GTTTATGTCACTAAATTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597  
Db 3541 GTTTATGTCACTAAATTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597

## RESULT 2

AA572657  
ID AAS72657 standard; cDNA; 3597 BP.

XX AC AAS72657;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #8461.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX XX (HYSB-) HYSBQ INC.

XX XX Drmanac RT, Liu C, Tang YT;

XX XX WPI; 2001-639362/73.

XX XX P-PSDB; ABG08470.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
diagnostics, forensics, gene mapping, identification of mutations  
responsible for genetic disorders or other traits and to assess  
biodiversity.

XX Claim 1; SEQ ID NO 8461; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and polypeptide (II)

XX sequences. (I) is useful as hybridisation probes, polymerase chain

XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping.

XX and in recombinant production of (II). The polynucleotides are also used

XX in diagnostics as expressed sequence tags for identifying expressed

XX genes. (I) is useful in gene therapy techniques to restore normal

XX activity of (II) or to treat disease states involving (II). (II) is

XX useful for generating antibodies against it, detecting or quantitating a

XX polypeptide in tissue, as molecular weight markers and as a food

XX supplement. (II) and its binding partners are useful in medical imaging

XX of sites expressing (II). (I) and (II) are useful for treating disorders

XX involving aberrant protein expression or biological activity. The

XX polypeptide and polynucleotide sequences have applications in

XX diagnostics, forensics, gene mapping, identification of mutations

XX responsible for genetic disorders or other traits to assess biodiversity

XX and to produce other types of data and products dependent on DNA and

XX amino acid sequences. AAS4137-AAS94564 represent novel human diagnostic

XX coding sequences of the invention. Note: the sequence data for this

XX patent did not appear in the printed specification, but was obtained in

XX electronic format directly from WIPO at

XX ftp.wipo.int/pub/published\_pct\_sequences

XX

XX Query Match 100.0%; Score 3597; DB 5; Length 3597;

XX Best Local Similarity 100.0%; Pred. No. 0;

XX Matches 3597; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGTGCGCCGAGCGCGGCACTGAGCGAAGCTGAGGAGGAGGAGCGCGGCGAGCGCA 60

Db 1 ATGTGCGCCGAGCGCGGCACTGAGCGAAGCTGAGGAGGAGGAGCGCGGCGAGCGCA 60

Qy 61 GCGCTCGCGGCGAGGAGCGCGGCGCACTGTGTCAACGGCGGTGCTGCAACCCCTGGC 120

Db 61 GCGCTCGCGGCGAGGAGCGCGGCGCACTGTGTCAACGGCGGTGCTGCAACCCCTGGC 120

Qy 121 ACCGGCGAGTGTCTGCGCGCGCTGAGCGGCAAGTGGCGGCAAGTGGCGGCGCTGCGGGGC 180

Db 121 ACCGGCGAGTGTCTGCGCGCGCTGAGCGGCAAGTGGCGGCAAGTGGCGGCGCTGCGGGGC 180

Qy 181 CGCTTCAGACTAATGAGTCTTCTGGGTTTGTGACAGATGGACCTGGAATATATAATAC 240

Db 181 CGCTTCAGACTAATGAGTCTTCTGGGTTTGTGACAGATGGACCTGGAATATATAATAC 240

Qy 241 AAAACGAAGTGCAGTGGCTCTATGAGGACAGCCAAATAGATATAGACTTCGTTTC 300

Db 241 AAAACGAAGTGCAGTGGCTCTATGAGGACAGCCAAATAGATATAGACTTCGTTTC 300

Qy 301 AATCAATTTGCTACAGAGTGTAGTGGGCAATTTATATGTTTATGATGGGACTCAAT 360

Db 301 AATCAATTTGCTACAGAGTGTAGTGGGCAATTTATATGTTTATGATGGGACTCAAT 360

Qy 361 TATGCAACGCTAGTGTGCAATTTAGTGGCTCATGTTCTCTGAGAGAGATGCAATGAG 420

Db 361 TATGCAACGCTAGTGTGCAATTTAGTGGCTCATGTTCTCTGAGAGAGATGCAATGAG 420

Qy 421 ACTGTCCCTGAGGTGTGGCAATCAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

Db 421 ACTGTCCCTGAGGTGTGGCAATCAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480

Qy 481 GCTTATTAATTTGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTCG 540

Db 481 GCTTATTAATTTGAGTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTCG 540

Qy 541 TCAGCGCGAGGAGTGTAGATCAGTAATAGGAGCGGCACTGTGATGTAATGATGTTCT 600

Db 541 TCAGCGCGAGGAGTGTAGATCAGTAATAGGAGCGGCACTGTGATGTAATGATGTTCT 600

Qy 601 GAAACCTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGCAACCTGTGTTTCT 660

Db 601 GAAACCTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGCAACCTGTGTTTCT 660

Qy 661 CATCGAGCATCTCAATCAAGTGTGAGGAGATGCTGCTGCTTCTCAGACTGCGCAG 720

Db 661 CATCGAGCATCTCAATCAAGTGTGAGGAGATGCTGCTGCTTCTCAGACTGCGCAG 720

Qy 721 GGTCTCGATGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 780

Db 721 GGTCTCGATGTTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 780

Qy 781 TCTAATCTTAAGCTCCCGAGGAGCATCTCAATAAGCTGTGCTCAATGGAACATATATG 840

Db 781 TCTAATCTTAAGCTCCCGAGGAGCATCTCAATAAGCTGTGCTCAATGGAACATATATG 840

Qy 841 GTTGTGAGGATATATGTTCAACCATCTCAGATTAATTAATTAATTAATTAATTAATTAAT 900

Db 841 GTTGTGAGGATATATGTTCAACCATCTCAGATTAATTAATTAATTAATTAATTAATTAAT 900

Qy 901 GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGTTGTAGATATG 960

Db 901 GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAATGTGTTGTAGATATG 960

Qy 961 CATTTCTTGGCATTAATCAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020

Db 961 CATTTCTTGGCATTAATCAAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1020

Qy 1021 GGGATGTGACCAATGAGTTGAGAGTGTTCACATTCATTAATGATCATGGGTGTTG 1080

Db 1021 GGGATGTGACCAATGAGTTGAGAGTGTTCACATTCATTAATGATCATGGGTGTTG 1080

Qy 1081 ACCCTTAAGGCAAGGAGGAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1140

Db 1081 ACCCTTAAGGCAAGGAGGAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGT 1140

Qy 1141 AAGATGCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1200

Db 1141 AAGATGCGCGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1200

Qy 1201 AGCAATGTGAGGAGTATGATTTGGATTAAGAACACATGGAGTATATTACACACCCAGG 1260

Db 1201 AGCAATGTGAGGAGTATGATTTGGATTAAGAACACATGGAGTATATTACACACCCAGG 1260

Qy 1261 GCCCTTGTGAGGAGGAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1320

Db 1261 GCCCTTGTGAGGAGGAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1320

Qy 1321 TACGTTTCAATGTTGCTACAGGCTTTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1380

Db 1321 TACGTTTCAATGTTGCTACAGGCTTTTCAAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1380

Qy 1381 TACGATATGATGAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1440

Db 1381 TACGATATGATGAGTATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1440

Qy 1441 TACTTGACACAGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1500

Db 1441 TACTTGACACAGCTGTGATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1500

Qy 1501 AATGACATCTATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1560

Db 1501 AATGACATCTATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1560

Qy 1561 ATTGCTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1620

Db 1561 ATTGCTGTGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1620

Qy 1621 TTGSCCATTCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1680

Db 1621 TTGSCCATTCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1680

Qy 1681 CTCCTCTCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1740

Db 1681 CTCCTCTCAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1740

```
QY 1741 GCCGCTTGTTCAGCAGGACCTGCTGATTCGCTGTGTGTGGAAACAAGGTCGTCTCAG 1800
DB 1741 GCCGCTTGTTCAGCAGGACCTGCTGATTCGCTGTGTGTGGAAACAAGGTCGTCTCAG 1800
QY 1801 TGTATCTCGTGGCGCTGGCAACTGATGAACAAGAAAGAAAGTTAAATCAGAAATGTTTT 1860
DB 1801 TGTATCTCGTGGCGCTGGCAACTGATGAACAAGAAAGAAAGTTAAATCAGAAATGTTTT 1860
QY 1861 TCCAAAGAACTCTTGACCAATGACAGATGTGACCAAGACACAGATTTGTTACAGTGCA 1920
DB 1861 TCCAAAGAACTCTTGACCAATGACAGATGTGACCAAGACACAGATTTGTTACAGTGCA 1920
QY 1921 GCCAAACCAATGACTGCCACTGGTGCATGACCAATGTTGTCGCCAGGAACACAGCTGC 1980
DB 1921 GCCAAACCAATGACTGCCACTGGTGCATGACCAATGTTGTCGCCAGGAACACAGCTGC 1980
QY 1981 TCAGAAGGCAGATCTCCATTTTATGATGAGAAATGGCCCAAGGAATACCCCTATGTAC 2040
DB 1981 TCAGAAGGCAGATCTCCATTTTATGATGAGAAATGGCCCAAGGAATACCCCTATGTAC 2040
QY 2041 TACTGTAAACAAGAACCCAGCTCAGAGCTGTGCCCTGAGACCAAGAACTGCCAGTGGAG 2100
DB 2041 TACTGTAAACAAGAACCCAGCTCAGAGCTGTGCCCTGAGACCAAGAACTGCCAGTGGAG 2100
QY 2101 CCCCGAATCAGGAGTGCATTTGCCCTGCCGAAATATCTGTGGCAATGGCTGGCATTTG 2160
DB 2101 CCCCGAATCAGGAGTGCATTTGCCCTGCCGAAATATCTGTGGCAATGGCTGGCATTTG 2160
QY 2161 GTTGGAACTCATGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2220
DB 2161 GTTGGAACTCATGTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 2220
QY 2221 TTCTGTAGGAACCAATGCCCCCTTTGGCTTCTTACAAACCAAGAAAGGTAGAAATTT 2280
DB 2221 TTCTGTAGGAACCAATGCCCCCTTTGGCTTCTTACAAACCAAGAAAGGTAGAAATTT 2280
QY 2281 GTCCTTAAGCAGCTGGCAATATGAGTCAATCTCAGAGATGTCAGAGTCACTTAAC 2340
DB 2281 GTCCTTAAGCAGCTGGCAATATGAGTCAATCTCAGAGATGTCAGAGTCACTTAAC 2340
QY 2341 CCATGGTGGGCTTCGGAAGATCAATGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
DB 2341 CCATGGTGGGCTTCGGAAGATCAATGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2400
QY 2401 TTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCAGTGAATCTGCTGGA 2460
DB 2401 TTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCAGTGAATCTGCTGGA 2460
QY 2461 ATTTTATCAGAACCCAGTACTCGGGACTGAGGCTGCAACCTGCATCAACCACTCAAT 2520
DB 2461 ATTTTATCAGAACCCAGTACTCGGGACTGAGGCTGCAACCTGCATCAACCACTCAAT 2520
QY 2521 GGTAGTGTCTGTGAAAGGCTGCAACCACTGCTAAGCAGTCCGGACACCATGTGCC 2580
DB 2521 GGTAGTGTCTGTGAAAGGCTGCAACCACTGCTAAGCAGTCCGGACACCATGTGCC 2580
QY 2581 TTGAGACAGCATGTGAGATTTGCAACCCAGGAGCTCTGAGTGCATGTGTGTCAGCAAC 2640
DB 2581 TTGAGACAGCATGTGAGATTTGCAACCCAGGAGCTCTGAGTGCATGTGTGTCAGCAAC 2640
QY 2641 ATGAAGCAGTGTGTGACTTCAATGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATG 2700
DB 2641 ATGAAGCAGTGTGTGACTTCAATGCTATGTGGCTCTTCCCTTTTGGCCAGTGTATG 2700
QY 2701 GAATGTGTATACGATGACACCTGCCCTCCCTGAAATTTGTTACGGCTACTGTACCTGTAGT 2760
DB 2701 GAATGTGTATACGATGACACCTGCCCTCCCTGAAATTTGTTACGGCTACTGTACCTGTAGT 2760
QY 2761 CATTTGTTGGAGAACCCAGGCTGTGCTGCTGTACTGATCCCGACCAATACTGGCAAGGG 2820
DB 2761 CATTTGTTGGAGAACCCAGGCTGTGCTGCTGTACTGATCCCGACCAATACTGGCAAGGG 2820
```

```
QY 2821 AAATGCATAGAGGTTCTCTATAAAGGACCACTGAGATGCTTCCGAGAGCCCTACAGGA 2880
DB 2821 AAATGCATAGAGGTTCTCTATAAAGGACCACTGAGATGCTTCCGAGAGCCCTACAGGA 2880
QY 2881 AATTTCTATCCACAGCCCTCTCTCAATTCACAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
DB 2881 AATTTCTATCCACAGCCCTCTCTCAATTCACAGCATGTGTCTAGAGGACAGCAGATACAAC 2940
QY 2941 TGGTCTTTTCAATCTACTGTCCAGCTTGCATATGCAAGGCAACAGTAAATGATCAATCAG 3000
DB 2941 TGGTCTTTTCAATCTACTGTCCAGCTTGCATATGCAAGGCAACAGTAAATGATCAATCAG 3000
QY 3001 AGCATCTGTGAGAAAGTGTGAGAACTGACACACAGGCAAGCACTGCGAGACCTGCATATCT 3060
DB 3001 AGCATCTGTGAGAAAGTGTGAGAACTGACACACAGGCAAGCACTGCGAGACCTGCATATCT 3060
QY 3061 GGCTCTTACGCTGATGCCACCAATGGAGGGAATATGTCAGCCATGCAATGAGGGGAC 3120
DB 3061 GGCTCTTACGCTGATGCCACCAATGGAGGGAATATGTCAGCCATGCAATGAGGGGAC 3120
QY 3121 CGCTCTCTGCAACCAACCAACGCGCAAGTCTTGCACCAACCAAGGGGCTCAAGGGG 3180
DB 3121 CGCTCTCTGCAACCAACCAACGCGCAAGTCTTGCACCAACCAAGGGGCTCAAGGGG 3180
QY 3181 GACGAGTCCAGCTATGTAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACA 3240
DB 3181 GACGAGTCCAGCTATGTAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACA 3240
QY 3241 TGTATTATATCTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGAT 3300
DB 3241 TGTATTATATCTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGAT 3300
QY 3301 CGCTATTACAGCTATCAATTTTGTGCTACTCTGACCAACCAACAGGGATTTGGAC 3360
DB 3301 CGCTATTACAGCTATCAATTTTGTGCTACTCTGACCAACCAACAGGGATTTGGAC 3360
QY 3361 ATGTTTCAATGCTTCAAGAAATTTCAACCTCAACATCAGTGGCTGCCAGTTCTCA 3420
DB 3361 ATGTTTCAATGCTTCAAGAAATTTCAACCTCAACATCAGTGGCTGCCAGTTCTCA 3420
QY 3421 GCTGGAACCCAGGCTGGAGAGAGATGCTGCTTCTTCAAAACCAACATTAAGGAGTAC 3480
DB 3421 GCTGGAACCCAGGCTGGAGAGAGATGCTGCTTCTTCAAAACCAACATTAAGGAGTAC 3480
QY 3481 AAGATAGTTTCTTAAATGAGAAAGTTTGAATTTCCGCAACCAACCAATATCCTTTCTTT 3540
DB 3481 AAGATAGTTTCTTAAATGAGAAAGTTTGAATTTCCGCAACCAACCAATATCCTTTCTTT 3540
QY 3541 GTTTATGTCAGTAATTTTCACTGCGCCCATCAAAATTCAGGTGCAAACTGACAAATGA 3597
DB 3541 GTTTATGTCAGTAATTTTCACTGCGCCCATCAAAATTCAGGTGCAAACTGACAAATGA 3597

RESULT 3
ADD70997
ID ADD70997 standard; DNA; 3597 BP.
XX
AC ADD70997;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human attractin gene SEQ ID NO:1.
XX
KW liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;
KW cytosolic; gene therapy; human; gene; ds.
XX
OS Homo sapiens.
XX
PN W02003061564-A2.
XX
PD 31-JUL-2003.
XX
PF 20-DEC-2002; 2002WO-US040718.
```



Db 1441 TACTTGACACACGCTGTGATAGTAGTGGAAACCATGCTGGTGTGTTGGGGAAACACACAC 1500  
Qy 1501 AATGACACATCTATGAGCCATGCGCAAAATGCTTCTTCACAGATTTCATGSCCTATGAC 1560  
Db 1501 AATGACACATCTATGAGCCATGCGCGCAAAATGCTTCTTCACAGATTTCATGSCCTATGAC 1560  
Qy 1561 ATTGGCTGTGACCGCTGGTCAAGTCTCCAGACCTGATTCACCAATGATGTCAACAGAT 1620  
Db 1561 ATTGGCTGTGACCGCTGGTCAAGTCTCCAGACCTGATTCACCAATGATGTCAACAGAT 1620  
Qy 1621 TTGGCCATTCCAGCAGCTTTACACAAACAGCACCATGTATGTGTTCGGTGTTCATAGT 1680  
Db 1621 TTGGCCATTCCAGCAGCTTTACACAAACAGCACCATGTATGTGTTCGGTGTTCATAGT 1680  
Qy 1681 CTCCTCTCAGCGACATCTCGGTATTCACCTCGAAACAGTGTGATCGGATCGGAGTAA 1740  
Db 1681 CTCCTCTCAGCGACATCTCGGTATTCACCTCGAAACAGTGTGATCGGATCGGAGTAA 1740  
Qy 1741 GCGGCTGTTTAGCAGCAGACCTGGTATTCGGGTGTGTGGAAACACAGGGTGTCTCAG 1800  
Db 1741 GCGGCTGTTTAGCAGCAGACCTGGTATTCGGGTGTGTGGAAACACAGGGTGTCTCAG 1800  
Qy 1801 TGTATCTGTGGGGCTGGCACTGATGAACAAAGAAAGAAAGTTAAATCAGAAATGTTTT 1860  
Db 1801 TGTATCTGTGGGGCTGGCACTGATGAACAAAGAAAGAAAGTTAAATCAGAAATGTTTT 1860  
Qy 1861 TCCAAAGAAACTCTTGACCATGACAGATGTGACGACACACAGATTGTTACAGCTGCACA 1920  
Db 1861 TCCAAAGAAACTCTTGACCATGACAGATGTGACGACACACAGATTGTTACAGCTGCACA 1920  
Qy 1921 GCCAACCAATCACTGCGCACTGGTCAATGACCAATTGTGTCGCCAGGAACACACAGCTGC 1980  
Db 1921 GCCAACCAATCACTGCGCACTGGTCAATGACCAATTGTGTCGCCAGGAACACACAGCTGC 1980  
Qy 1981 TCAGAGGCGCAGATCTCCATTTTAGGTATGAGAAATGGCCCAAGATTAACCCATATGTAC 2040  
Db 1981 TCAGAGGCGCAGATCTCCATTTTAGGTATGAGAAATGGCCCAAGATTAACCCATATGTAC 2040  
Qy 2041 TACTGTAAACAAAGAACACAGCTGCAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGGAG 2100  
Db 2041 TACTGTAAACAAAGAACACAGCTGCAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGGAG 2100  
Qy 2101 CCCCGAATCAGAGTGCATGCGCTGCCCGAAATATCTGTGGCAATTGGCTGGCAATTG 2160  
Db 2101 CCCCGAATCAGAGTGCATGCGCTGCCCGAAATATCTGTGGCAATTGGCTGGCAATTG 2160  
Qy 2161 GTTGGAAATCAATGTTTGAATAATCTACTGCCAAGGAGATTAACACAAATGCTAATG 2220  
Db 2161 GTTGGAAATCAATGTTTGAATAATCTACTGCCAAGGAGATTAACACAAATGCTAATG 2220  
Qy 2221 TTCTGTAGGAACCAATGCCCTTTTGGCTTCTCTTACAAACCAAGAAAGGTAGAAATTT 2280  
Db 2221 TTCTGTAGGAACCAATGCCCTTTTGGCTTCTCTTACAAACCAAGAAAGGTAGAAATTT 2280  
Qy 2281 GTCCCTTAAGCAGCTGGCAATTAATGCACTCATCTCAGAGCATGTCAGAGCTCACTTAACC 2340  
Db 2281 GTCCCTTAAGCAGCTGGCAATTAATGCACTCATCTCAGAGCATGTCAGAGCTCACTTAACC 2340  
Qy 2341 CCATGGGTGCGCCTTCGGAAGATCAATGTCTCCTACTGTGTGGGAAGATGTCCCA 2400  
Db 2341 CCATGGGTGCGCCTTCGGAAGATCAATGTCTCCTACTGTGTGGGAAGATGTCCCA 2400  
Qy 2401 TTTACAAATAGTTTACTACAGTGGATGCGCTGTGAGCCAGATGCTGGATCTGTGGA 2460  
Db 2401 TTTACAAATAGTTTACTACAGTGGATGCGCTGTGAGCCAGATGCTGGATCTGTGGA 2460  
Qy 2461 ATTTTATCAGAACCCAGTACTCGGGACTGGAAGCTGCAACCTGATCAACCCATCAAT 2520  
Db 2461 ATTTTATCAGAACCCAGTACTCGGGACTGGAAGCTGCAACCTGATCAACCCATCAAT 2520  
Qy 2521 GGTAGTGTCTGTGAAGGCTGCAAAACCAAGTGTAAAGCAGTGCAGCAATGTGCC 2580  
Db 2521 GGTAGTGTCTGTGAAGGCTGCAAAACCAAGTGTAAAGCAGTGCAGCAATGTGCC 2580

RESULT 4  
AAZ52275  
ID AAZ52275 standard; cDNA; 4068 BP.

Qy 2581 TTGAGGACAGCATGTGGAGATTGCAACGCGCAGCTCTGAGTGCATGTGTGAGCAAC 2640  
Db 2581 TTGAGGACAGCATGTGGAGATTGCAACGCGCAGCTCTGAGTGCATGTGTGAGCAAC 2640  
Qy 2641 ATGAAGCAGTGTGTGGACTCCAATGCTATGTGCGCTCTTCCCTTTTGGCCAGTGTATG 2700  
Db 2641 ATGAAGCAGTGTGTGGACTCCAATGCTATGTGCGCTCTTCCCTTTTGGCCAGTGTATG 2700  
Qy 2701 GAATGGTATACGATGAGCACCTGCGCCCTGAAATTTTTCAGGCTACTGTACCTGTAGT 2760  
Db 2701 GAATGGTATACGATGAGCACCTGCGCCCTGAAATTTTTCAGGCTACTGTACCTGTAGT 2760  
Qy 2761 CATTTGCTTTGGAGCAACAGGCTGTGGTGTGTACTGATCCAGCAATACTTGGCAAAAGG 2820  
Db 2761 CATTTGCTTTGGAGCAACAGGCTGTGGTGTGTACTGATCCAGCAATACTTGGCAAAAGG 2820  
Qy 2821 AAATGCATAGAGGGTTCCTATATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTTACAGGA 2880  
Db 2821 AAATGCATAGAGGGTTCCTATATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTTACAGGA 2880  
Qy 2881 AATTTCTATCCACAGCCCTGCTCAATTCACAGCATGTGTCTAGAGGACAGCAGATACAC 2940  
Db 2881 AATTTCTATCCACAGCCCTGCTCAATTCACAGCATGTGTCTAGAGGACAGCAGATACAC 2940  
Qy 2941 TGGTCTTTCAITTCATCTGTCAGCTTGCCAAATGCAACGGCCACAGTAAATGATCAATCAG 3000  
Db 2941 TGGTCTTTCAITTCATCTGTCAGCTTGCCAAATGCAACGGCCACAGTAAATGATCAATCAG 3000  
Qy 3001 AGCATCTGTGAGAGTGTGAGAACTGACCAACAGGCAAGCACTGCGAGACCTGCATATCT 3060  
Db 3001 AGCATCTGTGAGAGTGTGAGAACTGACCAACAGGCAAGCACTGCGAGACCTGCATATCT 3060  
Qy 3061 GCGTCTCTGTGCAACACCAACACACAGGCAAGTGTCTGACCCACCAAGGGGTCAAGGGG 3120  
Db 3061 GCGTCTCTGTGCAACACCAACACAGGCAAGTGTCTGACCCACCAAGGGGTCAAGGGG 3120  
Qy 3121 GCGTCTCTGTGCAACACCAACACAGGCAAGTGTCTGACCCACCAAGGGGTCAAGGGG 3180  
Db 3121 GCGTCTCTGTGCAACACCAACACAGGCAAGTGTCTGACCCACCAAGGGGTCAAGGGG 3180  
Qy 3181 GACGAGTGCAGCTATGTGAGGTAGAAATCGATACCAAGAAACCCCTCTCAGAGGAACA 3240  
Db 3181 GACGAGTGCAGCTATGTGAGGTAGAAATCGATACCAAGAAACCCCTCTCAGAGGAACA 3240  
Qy 3241 TGTATTATATCTCTCTTATTGACTATCAGTTTCACTTTTACTTCTATCCAGGAAGATGAT 3300  
Db 3241 TGTATTATATCTCTCTTATTGACTATCAGTTTCACTTTTACTTCTATCCAGGAAGATGAT 3300  
Qy 3301 CGCTATTACAGCTATCAATTTTGTGGTACTCTCTGACGAAACAAACAGGGATTTGGAC 3360  
Db 3301 CGCTATTACAGCTATCAATTTTGTGGTACTCTCTGACGAAACAAACAGGGATTTGGAC 3360  
Qy 3361 ATGTTTCATCAATGCTCCAGAAATTTCAACTCAACATCAGCTGGCTGCCAGTTTCTCA 3420  
Db 3361 ATGTTTCATCAATGCTCCAGAAATTTCAACTCAACATCAGCTGGCTGCCAGTTTCTCA 3420  
Qy 3421 GCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAACCAACATTAAGAGATAC 3480  
Db 3421 GCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAACCAACATTAAGAGATAC 3480  
Qy 3481 AAGAGTAGTTTCTCTAATGAGAAAGTTGATTTTTCGCAACCAACCAATATCATTCTTT 3540  
Db 3481 AAGAGTAGTTTCTCTAATGAGAAAGTTGATTTTTCGCAACCAACCAATATCATTCTTT 3540  
Qy 3541 GTTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597  
Db 3541 GTTTATGTCAGTAATTTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAACAATGA 3597



XX AC AAZ52275;  
 XX DT 18-JUL-2000 (first entry)  
 XX DE Human membrane attractin-1 cDNA.  
 XX KW Human; membrane attractin-1; immune response; macrophage; monocyte;  
 KW T cell; immunostimulant; immunosuppressed patient; cancer;  
 KW immunodeficiency syndrome; transplant; autoimmune disease; ss.  
 XX OS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 FT CDS 1..4068  
 FT /\*tag= a  
 FT /product= "Membrane attractin-1"  
 XX PN WO200015651-A1.  
 XX PD 23-MAR-2000.  
 XX PF 14-SEP-1999; 99WO-US020948.  
 XX PR 14-SEP-1998; 98US-0100137P.  
 XX PA (DAND ) DANA FARBER CANCER INST INC.  
 XX PI Duke-Cohan JS, Schlossman SF;  
 XX DR WPI; 2000-271373/23.  
 XX DR P-PSDB; AAY70690.  
 XX PT Isolated nucleic acids encoding human attractin polypeptides useful for  
 PT enhancing immune responses.  
 XX PS Claim 3; Fig 10; 120pp; English.  
 XX CC The patent discloses four forms of human attractin polypeptides which  
 CC enhance immune response by promoting macrophage and monocyte spreading in  
 CC the presence of T cells. These include soluble attractin-1 and -2 and  
 CC membrane attractin-1 and -2. These various forms of attractin are encoded  
 CC by alternatively spliced mRNA molecule transcribed from a single gene.  
 CC The present sequence is a cDNA encoding human membrane attractin-1.  
 CC Membrane attractin differs from soluble attractin in having a  
 CC transmembrane domain and a cytoplasmic domain at the C-terminal. This  
 CC sequence can be used to enhance immune response in immunosuppressed  
 CC patients such as those undergoing chemo- and radio-therapy treatment for  
 CC cancer or those suffering from common variable immunodeficiency syndrome.  
 CC The proteins may also be used to screen modulators (agonists and  
 CC antagonists) of immune response which may also be used to regulate immune  
 CC reactions. Attractin antibodies can be used to inhibit immune response in  
 CC transplant recipients or patients afflicted with autoimmune disease  
 XX Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;  
 XX  
 XX Query Match 98.9%; Score 3556.4; DB 3; Length 4068;  
 XX Best Local Similarity 99.9%; Pred. No. 0;  
 XX Matches 3579; Conservative 0; Mismatches 1; Indels 2; Gaps 2;  
 QY 1 ATGTGGCCGCGAGCGGCGCACTGAGCAAGGTGAGGAGGAGCGCGCGGCGGCA 60  
 DB 1 ATGTGGCCGCGAGCGGCGCACTGAGCAAGGTGAGGAGGAGCGCGCGGCGGCA 60  
 QY 61 GCGTCGCGGCGAGAGCGGCGCGCACTGTGTCAACGGCGGTGCTGCAACCCCTGGC 120  
 DB 61 GCGTCGCGGCGAGAGCGGCGCGCACTGTGTCAACGGCGGTGCTGCAACCCCTGGC 120  
 QY 121 ACCGGCCAGTCGTCGCGCGCGGTGGTGGCGGAGCAATGCCAGCACTCGGGGGC 180  
 DB 121 ACCGGCCAGTCGTCGCGCGCGGTGGTGGCGGAGCAATGCCAGCACTCGGGGGC 180  
 QY 181 CGCTTCAGACTACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAATAC 240

DB 181 CGCTTCAGACTACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTATAATAC 240  
 QY 241 AAAACGAAGTGCACGTGGCTCAATTGAAGGACAGCAATAGATAATAGACATTCGTTTC 300  
 DB 241 AAAACGAAGTGCACGTGGCTCAATTGAAGGACAGCAATAGATAATAGACATTCGTTTC 300  
 QY 301 AATCATTTTGTACAGAGTGTAGTTGGGACCATTTATATATTTATGATGGGACTCAATT 360  
 DB 301 AATCATTTTGTACAGAGTGTAGTTGGGACCATTTATATATTTATGATGGGACTCAATT 360  
 QY 361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCCTCAATTGCTGAGAGAGATGGCAATGAG 420  
 DB 361 TATGCACCGCTAGTTGCTGCATTTAGTGGCCCTCAATTGCTGAGAGAGATGGCAATGAG 420  
 QY 421 ACTGTCCCTGAGGTTGTGCCACATCAGGTTATGCTTGTGCTGCTGCTGCTGCTGCTGCT 480  
 DB 421 ACTGTCCCTGAGGTTGTGCCACATCAGGTTATGCTTGTGCTGCTGCTGCTGCTGCTGCT 480  
 QY 481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTGTGATATGCTCCAAATAACTGC 540  
 DB 481 GCTTATAATTTGACTGGATTTAATATTACTTACAGTTTGTGATATGCTCCAAATAACTGC 540  
 QY 541 TCAGGCCGAGGAGAGTGTAGATCAGTAATAGCAGGATCTGTTGTAATGTGAATGTTCT 600  
 DB 541 TCAGGCCGAGGAGAGTGTAGATCAGTAATAGCAGGATCTGTTGTAATGTGAATGTTCT 600  
 QY 601 GAAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660  
 DB 601 GAAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCT 660  
 QY 661 CATCGAGGCATCTCAATTCAGATGTGACAGGATGCTCCTGCTCTCAGACTGCGAG 720  
 DB 661 CATCGAGGCATCTCAATTCAGATGTGACAGGATGCTCCTGCTCTCAGACTGCGAG 720  
 QY 721 GGTCTCGATGTTCAAGTTCCTGACAGCACTCACTTTCGAGTCCAGAGGAAATAT 780  
 DB 721 GGTCTCGATGTTCAAGTTCCTGACAGCACTCACTTTCGAGTCCAGAGGAAATAT 780  
 QY 781 TCTAACTTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGTCATGGAACAACTATGTGG 840  
 DB 781 TCTAACTTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGTCATGGAACAACTATGTGG 840  
 QY 841 GTTGTGAGAGATATGTTTCAACCACTCAGATTTACACATGTTTACAGGATGACCTT 900  
 DB 841 GTTGTGAGAGATATGTTTCAACCACTCAGATTTACACATGTTTACAGGATGACCTT 900  
 QY 901 GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAACTGTGGTTTGTAGATATGGT 960  
 DB 901 GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACAACTGTGGTTTGTAGATATGGT 960  
 QY 961 CATTCTTTGGCATTTATACAGGATAAATTTACATGTTATGAGGAAATTTGATCAACT 1020  
 DB 961 CATTCTTTGGCATTTATACAGGATAAATTTACATGTTATGAGGAAATTTGATCAACT 1020  
 QY 1021 GGGAAATGTGACCAATGAGTGTGAGAGTTTTTCACATTCATATAATGAGTGTGTTG 1080  
 DB 1021 GGGAAATGTGACCAATGAGTGTGAGAGTTTTTCACATTCATATAATGAGTGTGTTG 1080  
 QY 1081 ACCCTTAAGGCAAGGAGCAGTATGCGAGTGTGGGCACTCTGACACATTTGTACACTG 1140  
 DB 1081 ACCCTTAAGGCAAGGAGCAGTATGCGAGTGTGGGCACTCTGACACATTTGTACACTG 1140  
 QY 1141 AAGAAATGCCAGTGTGCTCATCTTGTGCTCACTGCCCTCTCTATGGAATATA 1200  
 DB 1141 AAGAAATGCCAGTGTGCTCATCTTGTGCTCACTGCCCTCTCTATGGAATATA 1200  
 QY 1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260  
 DB 1201 AGCAATGTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGT 1260  
 QY 1261 GCCCTTTGTCAGGGGGTTACGGCCATAGCAGTGTGTTACGACCATAGGACCCAGGGCCCTA 1320

Db 1261 GCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTGTTTACGACCATAGGACCAGGGCCCTA 1320  
Qy 1321 TAGCTTCATGGTGGCTCAAGGCTTTTCAGTGGCCAAATAGTACCGGCTTCGAGATGATCTC 1380  
Db 1321 TAGCTTCATGGTGGCTCAAGGCTTTTCAGTGGCCAAATAGTACCGGCTTCGAGATGATCTC 1380  
Qy 1381 TACCGATATGATGTGATATCCAGATGCCAGATGTGACCAATCTTAAAGACAGCCGATTTTCGGT 1440  
Db 1381 TACCGATATGATGTGATATCCAGATGTGACCAATCTTAAAGACAGCCGATTTTCGGT 1440  
Qy 1441 TACTTGCACACAGCTGTGATAGTACGTGGAAACCATCTGGTGTGTTGGGGGAAACACACAC 1500  
Db 1441 TACTTGCACACAGCTGTGATAGTACGTGGAAACCATCTGGTGTGTTGGGGGAAACACACAC 1500  
Qy 1501 AATGACACATCTATAGCCATAGCCATGCGCCAAATGCTTCTCTTCAGATTCATGCGCTATGAC 1560  
Db 1501 AATGACACATCTATAGCCATAGCCATGCGCCAAATGCTTCTCTTCAGATTCATGCGCTATGAC 1560  
Qy 1561 ATTGCCCTGTGACCGCTGGTCAAGTGTCCAGACCTGATCTCCACCATGATGTCACACAGA 1619  
Db 1561 ATTGCCCTGTGACCGCTGGTCAAGTGTCCAGACCTGATCTCCACCATGATGTCACACAGA 1620  
Qy 1620 TTTGGCCATTCAGAGTCTTACACAAACAGCACCATGTATGTGTTCCGGTGGTTTCAATAG 1679  
Db 1621 TTTGGCCATTCAGAGTCTTACACAAACAGCACCATGTATGTGTTCCGGTGGTTTCAATAG 1679  
Qy 1680 TCTCCTCTCAGCGACATCTCTGTATTCACCTCGGAACAGTGTGATCGGCATCGGAGTGA 1739  
Db 1680 TCTCCTCTCAGCGACATCTCTGTATTCACCTCGGAACAGTGTGATCGGCATCGGAGTGA 1739  
Qy 1740 AGCGGTGTTTACAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGTCGTCTCA 1799  
Db 1740 AGCGGTGTTTACAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGTCGTCTCA 1799  
Qy 1800 GTGTATCTGTGGGCGCTGCACTGATGAACAGAGAAAGTTAAATCAGATGTTT 1859  
Db 1800 GTGTATCTGTGGGCGCTGCACTGATGAACAGAGAAAGTTAAATCAGATGTTT 1859  
Qy 1860 TTTCCAAAAGAACCTTTGACCATGACAGATGTGACACAGCACACAGATTTGTACAGCTGAC 1919  
Db 1860 TTTCCAAAAGAACCTTTGACCATGACAGATGTGACACAGCACACAGATTTGTACAGCTGAC 1919  
Qy 1920 AGCCAAACCAATGATCGGCACCTGGTGCAATGACCATGTGTCGCCAGGAACACAGCTG 1979  
Db 1920 AGCCAAACCAATGATCGGCACCTGGTGCAATGACCATGTGTCGCCAGGAACACAGCTG 1979  
Qy 1980 CTGAGAGGCCAGATCTCAATTTTATAGTATGAGATTCGCCAAGGATAACCTATGTA 2039  
Db 1980 CTGAGAGGCCAGATCTCAATTTTATAGTATGAGATTCGCCAAGGATAACCTATGTA 2039  
Qy 2040 CTAATGTAAACAGACAGCTGACAGGAGCTGTCCTGGACACAGAACTGCCAGTGGGA 2099  
Db 2040 CTAATGTAAACAGACAGCTGACAGGAGCTGTCCTGGACACAGAACTGCCAGTGGGA 2099  
Qy 2100 GCCCGGAATCAGAGTGCATTCGCTGCCGGAATAATCTGTGGCATTTGGCTGTCATTT 2159  
Db 2100 GCCCGGAATCAGAGTGCATTCGCTGCCGGAATAATCTGTGGCATTTGGCTGTCATTT 2159  
Qy 2160 GGTGTGAAACTCATGTTGAAATTTACTACTGCGAAGGAGATTTATGACAAATGCTAAAT 2219  
Db 2160 GGTGTGAAACTCATGTTGAAATTTACTACTGCGAAGGAGATTTATGACAAATGCTAAAT 2219  
Qy 2220 GTTCTGTAGAAACCAATGCGCTTTTGGCTTCTTACAAACAGAGAGGTAGAAAT 2279  
Db 2220 GTTCTGTAGAAACCAATGCGCTTTTGGCTTCTTACAAACAGAGAGGTAGAAAT 2279  
Qy 2280 TGTCTTTAAGCAGCTGCAATATGAGTCACTCAGAGCATGTCCAAAGCTCACCTTAAC 2339  
Db 2280 TGTCTTTAAGCAGCTGCAATATGAGTCACTCAGAGCATGTCCAAAGCTCACCTTAAC 2339  
Qy 2340 CCCATGGTGGGCTTCGGAGATCAATGTGTCTTACTGGTGTGCGGAAGATATGTCCCC 2399  
Db 2340 CCCATGGTGGGCTTCGGAGATCAATGTGTCTTACTGGTGTGCGGAAGATATGTCCCC 2399

Qy 2400 ATTTTCAAAATAGTTTACTACAGTGGATGCCCTGTGAGCCCAAGTGAATGCTGATTTCTGTGG 2459  
Db 2400 ATTTTCAAAATAGTTTACTACAGTGGATGCCCTGTGAGCCCAAGTGAATGCTGATTTCTGTGG 2459  
Qy 2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCACTCAACCCACTCAA 2519  
Db 2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCACTCAACCCACTCAA 2519  
Qy 2520 TGGTAGTGTCTGTGAAAGGCGCTGCAAAACCAAGTGTAAAGCAGTGCCTGGGACACCATGTGC 2579  
Db 2520 TGGTAGTGTCTGTGAAAGGCGCTGCAAAACCAAGTGTAAAGCAGTGCCTGGGACACCATGTGC 2579  
Qy 2580 CTTGAGGACACATGTGGAGATTGACACAGCGGCGAGCTCTGAGTGCATGTGGTGCAGCAA 2639  
Db 2580 CTTGAGGACACATGTGGAGATTGACACAGCGGCGAGCTCTGAGTGCATGTGGTGCAGCAA 2639  
Qy 2640 CATGAACAGTGTGTGAGACTCCAAATGCTATGTGGCCTCTTCCCTTTTGGCCAGTGTAT 2699  
Db 2640 CATGAACAGTGTGTGAGACTCCAAATGCTATGTGGCCTCTTCCCTTTTGGCCAGTGTAT 2699  
Qy 2700 GGAATGTATACGATGAGACCTGCCCTGAAAATGTTTCAGGCTACTGTACTGTAG 2759  
Db 2700 GGAATGTATACGATGAGACCTGCCCTGAAAATGTTTTCAGGCTACTGTACTGTAG 2759  
Qy 2760 TCATTGCTTGGAGCAACCAAGCTGTGGCTGTACTGATCCAGCAATACTGCAAGG 2819  
Db 2760 TCATTGCTTGGAGCAACCAAGCTGTGGCTGTACTGATCCAGCAATACTGCAAGG 2819  
Qy 2820 GAAATGCATAGAGGTTTCTTATAAGGACGATGCAAGATGCTTCGAAGCCCTACAGG 2879  
Db 2820 GAAATGCATAGAGGTTTCTTATAAGGACGATGCAAGATGCTTCGAAGCCCTACAGG 2879  
Qy 2880 AAATTTTATCCACAGCCCTGCTCAATTCAGCATGTCTAGAGGACAGCATACAA 2939  
Db 2880 AAATTTTATCCACAGCCCTGCTCAATTCAGCATGTCTAGAGGACAGCATACAA 2939  
Qy 2940 CTGCTTTTCAATCACTGTCCAGTTCGAATGCAACGGGCCACAGTAAATGCATCAATCA 2999  
Db 2940 CTGCTTTTCAATCACTGTCCAGTTCGAATGCAACGGGCCACAGTAAATGCATCAATCA 2999  
Qy 3000 GAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAAGCACCTGCCAGACTGCATATC 3059  
Db 3000 GAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAAGCACCTGCCAGACTGCATATC 3059  
Qy 3060 TGGCTTTCTACGGTGTATCCCAATGGAGGAAATGTCAAGTGCATGCAATGGCA 3119  
Db 3060 TGGCTTTCTACGGTGTATCCCAATGGAGGAAATGTCAAGTGCATGCAATGGCA 3119  
Qy 3120 CGCGTCTGTGCAACACCAACAGCGGCAAGTGTCTGCAACCAAGGCGTCAAGGG 3179  
Db 3120 CGCGTCTGTGCAACACCAACAGCGGCAAGTGTCTGCAACCAAGGCGTCAAGGG 3179  
Qy 3180 GGACAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCTCTCAGAGAAC 3239  
Db 3180 GGACAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCTCTCAGAGAAC 3239  
Qy 3240 ATGTTATTATCTCTTTATTGATCTATCAGTTCACTTGTAGTCTATCCAGGAGATGA 3299  
Db 3240 ATGTTATTATCTCTTTATTGATCTATCAGTTCACTTGTAGTCTATCCAGGAGATGA 3299  
Qy 3300 TCGCTATTATACAGCTCAATTTTGTGGCTACTCTGACGAAACAAACAGGATTTTGA 3359  
Db 3300 TCGCTATTATACAGCTCAATTTTGTGGCTACTCTGACGAAACAAACAGGATTTTGA 3359  
Qy 3360 CATGTTCAATGCTCCAGAAATTTCAACCTCAACATCACTCCGGGCTGCAGTTTCTC 3419  
Db 3360 CATGTTCAATGCTCCAGAAATTTCAACCTCAACATCACTCCGGGCTGCAGTTTCTC 3419  
Qy 3420 AGCTGGAACCCAGCTGGAGAGATGCTGTTGTTTCAAAAACCAATTAAGGAGTA 3479  
Db 3420 AGCTGGAACCCAGCTGGAGAGATGCTGTTGTTTCAAAAACCAATTAAGGAGTA 3479

QY 3480 CAAAGATAGTTTCTTAATGAGAAGTTTGATTTTGGCAACCAACCCAAATATCATCTTTCTT 3539  
 DB |||||  
 DB 3480 CAAAGATAGTTTCTTAATGAGAAGTTTGATTTTGGCAACCAACCCAAATATCATCTTTCTT 3539  
 QY 3540 TGTATTGTCAGTAATTTTCACTGGCCCATCAAAATTCAGGT 3581  
 DB |||||  
 DB 3540 TGTATTGTCAGTAATTTTCACTGGCCCATCAAAATTCAGAT 3581

RESULT 5  
 AAZ91921  
 ID AAZ91921 standard; cDNA; 4072 BP.  
 XX  
 AC AAZ91921;  
 XX  
 DT  
 XX 08-JUN-2000 (first entry)  
 DE Human mahogany protein coding sequence #3.  
 XX  
 KW Mahogany gene; human; mg gene; regulatory defect: gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antioesity; antianorexic; anticachexic; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200005373-A2.  
 XX  
 PD 03-FEB-2000.  
 XX  
 XX 21-JUL-1999; 99WO-US016484.  
 XX  
 XX 21-JUL-1998; 98US-0093630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 99US-00245041.  
 XX  
 XX (MILL-) MILLENIUM PHARM INC.  
 PA  
 XX Moore K, Nagle DL;  
 PI  
 XX WPI; 2000-195103/17.  
 DR P-PSDB; AAY81808.  
 DR  
 XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.  
 XX  
 XX Claim 1; Fig 19a; 188pp; English.  
 XX  
 CC This sequence represents a human mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (II) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (II) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (II)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 CC  
 XX Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;  
 SQ

Query Match 97.1%; Score 3494.4; DB 3; Length 4072;  
 Best Local Similarity 98.8%; Pred. No. 0;  
 Matches 3541; Conservative 0; Mismatches 41; Indels 2; Gaps 2;

QY 15 GCGGCAACTCAGGCAAGGCTGAGAGGAGGAGCGCGCGGACGCGCTCGGGGCGAG 74  
 DB |||||  
 DB 234 GGAGGCGAGGCGCGCTGCGGTGGCGGGCGGGTGTCTCGGCTCGGCGCGCGAGGCCAA 293  
 QY 75 GAGCGGCGGGCGCGCACTGTGTCAACGGCGGTGCGTGCACCCCTGGCAACCGGCGAGTGCCT 134  
 DB |||||  
 DB 294 GGAATGTGACCGGCGGTGTGCAACGGCGGTGCGTGCACCCCTGGCAACCGGCGAGTGCCT 353  
 QY 135 CTGCCCCCGCGCTGGGTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAAC 194  
 DB |||||  
 DB 354 CTGCCCCCGCGCTGGGTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAAC 413  
 QY 195 TGGATCTTCTGGGTTTGTGACAGATGCACCTGGAAATATATAAATACAAAAAGAGTGCAC 254  
 DB |||||  
 DB 414 TGGATCTTCTGGGTTTGTGACAGATGCACCTGGAAATATATAAATACAAAAAGAGTGCAC 473  
 QY 255 GTGCTCATTCGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTAC 314  
 DB |||||  
 DB 474 GTGCTCATTCGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTAC 533  
 QY 315 AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGAGCTCAATTTATGCACCGCTAGT 374  
 DB |||||  
 DB 534 AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGAGCTCAATTTATGCACCGCTAGT 593  
 QY 375 TGTGCAATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434  
 DB |||||  
 DB 594 TGTGCAATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653  
 QY 435 TGTGCGCACATCAGGTTATGCTTGTGCAATTTTATGATGCTGCTCTATATTTGAC 494  
 DB |||||  
 DB 654 TGTGCGCACATCAGGTTATGCTTGTGCAATTTTATGATGCTGCTCTATATTTGAC 713  
 QY 495 TGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 554  
 DB |||||  
 DB 714 TGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 773  
 QY 555 GTGTAAGATCAGTAATAGCAGGATACCTGTTGAATGTGAATGTTCTGAAATCGGAAGG 614  
 DB |||||  
 DB 774 GTGTAAGATCAGTAATAGCAGGATACCTGTTGAATGTGAATGTTCTGAAATCGGAAGG 833  
 QY 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCTCATCGAGGCATCTG 674  
 DB |||||  
 DB 834 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCTCATCGAGGCATCTG 893  
 QY 675 CAATTCAGATGATGTGAGAGATGCTCTGCTCTCAGACTGGCAGGCTCTGATGTTTC 734  
 DB |||||  
 DB 894 CAATTCAGATGATGTGAGAGATGCTCTGCTCTCAGACTGGCAGGCTCTGATGTTTC 953  
 QY 735 AGTTCTGTACCACTACCACTCAATTTGGACCTGAGAGGAATATTCTAACTTTAAAGCT 794  
 DB |||||  
 DB 954 AGTTCTGTACCACTACCACTCAATTTGGACCTGAGAGGAATATTCTAACTTTAAAGCT 1013  
 QY 795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAACAAATATGTTGGTGTGTGGAGGATA 854  
 DB |||||  
 DB 1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAACAAATATGTTGGTGTGTGGAGGATA 1073  
 QY 855 TATGTTTCAACCACTCAGATTTATACATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 914  
 DB |||||  
 DB 1074 TATGTTTCAACCACTCAGATTTATACATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG 1133  
 QY 915 GCTTCCCACTAAACCGTTCTGTGAACAATGCTGTTTGTAGATPATGCTCATTTTGGCAAT 974  
 DB |||||  
 DB 1134 GCTTCCCACTAAACCGTTCTGTGAACAATGCTGTTTGTAGATPATGCTCATTTTGGCAAT 1193  
 QY 975 ATACAAGATATAATTTACATGTTATGAGGAAATATGATTTCACTGGGAATGTGACCA 1034  
 DB |||||  
 DB 1194 ATACAAGATATAATTTACATGTTATGAGGAAATATGATTTCACTGGGAATGTGACCA 1253  
 QY 1035 TGAGTTGAGAGTTTTTTCATTCATTAATAGCTCATGGGTGTTGTGACCCCTAAGGCAAA 1094  
 DB |||||  
 DB 1254 TGAGTTGAGAGTTTTTTCATTCATTAATAGCTCATGGGTGTTGTGACCCCTAAGGCAAA 1313  
 QY 1095 GGAGCAGTATGAGTGGTGGGCTCTGTCACACATTTGTACACTGAAGATGCGCGAGT 1154

1314 GAGCAGATGCGAGTGGTGGGCACTGTGCACACATTTGTACACTGAAGATGGCGAGT 1373  
1155 GGTGATGCTGGTCACTCTTTGGTCACTGCCCTCTATGGATATATAAGCAATGTGCAGCA 1214  
1374 GGTGATGCTGGTCACTCTTTGGTCACTGCCCTCTATGGATATATAAGCAATGTGCAGCA 1433  
1215 ATATGATTTGGATAAGAACACATGGAGATATATTACACACCCAGGGTGGCTTTGTGCAAGG 1274  
1434 ATATGATTTGGATAAGAACACATGGAGATATATTACACACCCAGGGTGGCTTTGTGCAAGG 1493  
1275 GGGTTACGGCCATAGCAGTGTGTTACGACCATAGACACCGGGCCCTATAGTTCATCGGTG 1334  
1494 GGGTTACGGCCATAGCAGTGTGTTACGACCATAGACACCGGGCCCTATAGTTCATCGGTG 1553  
1335 CTACAGGCTTTTCACTGCCAATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGT 1394  
1554 CTACAGGCTTTTCACTGCCAATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGT 1613  
1395 GGATACCCAGATGGACCACTCTTAAGGACACCGGATTTTCCGTTACTTGCACACAGC 1454  
1614 GGATACCCAGATGGACCACTCTTAAGGACACCGGATTTTCCGTTACTTGCACACAGC 1673  
1455 TGTGATAGTCAAGTGGAAACCATGCTGTGTTTGGGGGAAACACACAAATGACACATCTAT 1514  
1674 TGTGATAGTCAAGTGGAAACCATGCTGTGTTTGGGGGAAACACACAAATGACACATCTAT 1733  
1515 GAGCCATGGCGCCAAATGCTTCTTCAGATTTTCATGGCCATGACATTTGCCGTGACCG 1574  
1734 GAGCCATGGCGCCAAATGCTTCTTCAGATTTTCATGGCCATGACATTTGCCGTGACCG 1793  
1575 CTGGTCACTGCTTCCAGACCTGAT - TCCACCATGATGTCACACAGATTTGGCCATTCCAG 1633  
1794 CTGGTCACTGCTTCCAGACCTGATCTCCACCATGATGTCACACAGATTTGGCCATT - CAG 1852  
1634 CAGTCTTACACACAGACCATGATGTGTTGGTGGTTTCAATAGTCTCTCCCTCAGCG 1693  
1853 CAGTCTTACACACAGACCATGATGTGTTGGTGGTTTCAATAGTCTCTCCCTCAGCG 1912  
1694 ACATCTGTTATTCACCTCGAACAGTGTGATCGGCATCGGATGAGCGGCTGTTTAG 1753  
1913 ACATCTGTTATTCACCTCGAACAGTGTGATCGGCATCGGATGAGCGGCTGTTTAG 1972  
1754 CAGCAGACCTGGTATTCGGTGTGTGTGGAACACACAGGCTGCTCAGTGTATCTCGTGGG 1813  
1973 CAGCAGACCTGGTATTCGGTGTGTGTGGAACACACAGGCTGCTCAGTGTATCTCGTGGG 2032  
1814 CGCTGGCACTGATGAACAAGAAAGATTTAAATCAGATGTTTTTCCAAAGAACTC 1873  
2033 CGCTGGCACTGATGAACAAGAAAGATTTAAATCAGATGTTTTTCCAAAGAACTC 2092  
1874 TTGACCATGACATGTGACACGACACACAGATTTTACAGTGCACAGCCACACCAATG 1933  
2093 TTGACCATGACATGTGACACGACACACAGATTTTACAGTGCACAGCCACACCAATG 2152  
1934 ACTGGCACTGGTGAATGACCATTTGTGCCAGGAACACACAGTCTCAGAGGCCAGA 1993  
2153 ACTGGCACTGGTGAATGACCATTTGTGCCAGGAACACACAGTCTCAGAGGCCAGA 2212  
1994 TCTCCATTTTGTAGTATGAAATTTGCCCAAGGATTAACCTATGTACTCTGTACAGA 2053  
2213 TCTCCATTTTGTAGTATGAAATTTGCCCAAGGATTAACCTATGTACTCTGTACAGA 2272  
2054 AGACCACTGACGAGTGTGCCCTTGACACAGAACTGCCAGTGGAGCCCGGAATCAGG 2113  
2273 AGACCACTGACGAGTGTGCCCTTGACACAGAACTGCCAGTGGAGCCCGGAATCAGG 2332  
2114 AGTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCAATTTGGTGGAACTCAT 2173  
2333 AGTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGTGGCAATTTGGTGGAACTCAT 2392  
2174 GTTTGAAATTACTACTGCCAGGAAATTTATCACAATGCTAAATGTTTCTGTAGGAACC 2233

2393 GTTTGAAATTACTACTGCCAGGAGAAATTAAGCAATGCTAAATGTTCTGTAGGAACC 2452  
2234 ACAATGCCCTTTTGGCTTCTCTTACCAACCCAGAGAAGGTAGAATTTGCTCTTAAGCAGC 2293  
2453 ACAATGCCCTTTTGGCTTCTCTTACCAACCCAGAGAAGGTAGAATTTGCTCTTAAGCAGC 2512  
2294 TCGGAATTAATGAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGGCC 2353  
2513 TCGGAATTAATGAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGGCC 2572  
2354 TTTGGAAGATCAATGTGTCTTACTGGTCTGGGAAGATATGTCCTCAATTTTACAAATAGTT 2413  
2573 TTTGGAAGATCAATGTGTCTTACTGGTCTGGGAAGATATGTCCTCAATTTTACAAATAGTT 2632  
2414 TACTACAGTGGATGCGCTCTGAGCCCACTGATGTGGATTTCTGTGGAATTTTACAGAAC 2473  
2633 TACTACAGTGGATGCGCTCTGAGCCCACTGATGTGGATTTCTGTGGAATTTTACAGAAC 2692  
2474 CCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2533  
2693 CCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTG 2752  
2534 AAAGGCCCTGCAAAACACACAGTCTTAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCAT 2593  
2753 AAAGGCCCTGCAAAACACACAGTCTTAAGCAGTCCCGGACACCATGTGCTTTGAGGACAGCAT 2812  
2594 GTGGAGATTTGCACACCGGCACTGTGATGTCATGTGGTGCAGCAACATGAGCAGTGTG 2653  
2813 GTGGAGATTTGCACACCGGCACTGTGATGTCATGTGGTGCAGCAACATGAGCAGTGTG 2872  
2654 TGGACTCCAATGCTATGTGGCCCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2713  
2873 TGGACTCCAATGCTATGTGGCCCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGA 2932  
2714 TGAGCACCTGCCCCCTGAAAATTTTTCAGGCTACTGTACCTGTAGTCAITTTGTTGAGC 2773  
2933 TGAGCACCTGCCCCCTGAAAATTTTTCAGGCTACTGTACCTGTAGTCAITTTGTTGAGC 2992  
2774 AACAGGCTGTGGCTGGTGTACTGATCCAGCAATCTGCGCAAGGGAATGCAATAGAGG 2833  
2993 AACAGGCTGTGGCTGGTGTACTGATCCAGCAATCTGCGCAAGGGAATGCAATAGAGG 3052  
2834 GTTCTATAAAGGACCACTGGAAGTGCCTTCGCAAGCCCTTACAGGAAATTTCTATCCAC 2893  
3053 GTTCTATAAAGGACCACTGGAAGTGCCTTCGCAAGCCCTTACAGGAAATTTCTATCCAC 3112  
2894 AGCCCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCATTC 2953  
3113 AGCCCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCATTC 3172  
2954 ACTGTCCAGCTTGCCAATGCAACCGCCACAGTAATGATCAATCAGAGCATCTGTGAGA 3013  
3173 ACTGTCCAGCTTGCCAATGCAACCGCCACAGTAATGATCAATCAGAGCATCTGTGAGA 3232  
3014 AGTGTGAAACCTGACACACAGGCAAGCACTGCGAGACCTGCAATCTGGCTTTACGGTG 3073  
3233 AGTGTGAAACCTGACACACAGGCAAGCACTGCGAGACCTGCAATCTGGCTTTACGGTG 3292  
3074 ATCCCAACCAATGGAGGAAATGTGAGCCATGCAATGCAATGGGCAAGCTCTGTGCA 3133  
3293 ATCCCAACCAATGGAGGAAATGTGAGCCATGCAATGCAATGGGCAAGCTCTGTGCA 3352  
3134 ACACCAACACCGGCAAGTGTCTTTCGACCAACCAAGGGGCTCAAGGGGACAGGTGCCAGC 3193  
3353 ACACCAACACCGGCAAGTGTCTTTCGACCAACCAAGGGGCTCAAGGGGACAGGTGCCAGC 3412  
3194 TATGTGAGGTAGAAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTC 3253  
3413 TATGTGAGGTAGAAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTC 3472  
3254 TTTCTATTGACTATCAGTTTCACTTTTGTAGTCTATCCAGGAAGATGATCGCTATTACACAG 3313  
3473 TTTCTATTGACTATCAGTTTCACTTTTGTAGTCTATCCAGGAAGATGATCGCTATTACACAG 3532



Db 1170 TGTAGATAGGTCATCTTTGGCATTATACAGGATATAATTTACATGATGGAGAAA 1229  
Qy 1008 AATTGATTAACATGGGAATGTGACCAATGATGATGAGATTTTTCACATTCATAATGAGTC 1067  
Db 1230 AATTGATCCAACTGGGAATGTGACCAATGATGATGAGATTTTTCACATTCATAATGAGTC 1289  
Qy 1068 ATGGGTGTTCTGTACCCCTTAAGCCAAAGAGGAGTATGCAATGGTTGGGCACTCTGCACA 1127  
Db 1290 ATGGGTGTTCTGTACCCCTTAAGCCAAAGAGGAGTATGCAATGGTTGGGCACTCTGCACA 1349  
Qy 1128 CATTTGTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCACTTTTGGTCACTGGCCCTCT 1187  
Db 1350 CATTTGTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCACTTTTGGTCACTGGCCCTCT 1409  
Qy 1188 CTATGGATATATAGCAATGTGAGGATATGATTTGGATAGAAACACATGAGTATATT 1247  
Db 1410 CTATGGATATATAGCAATGTGAGGATATGATTTGGATAGAAACACATGAGTATATT 1469  
Qy 1248 ACACACCCAGGGTGCCTCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1307  
Db 1470 ACACACCCAGGGTGCCTCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAG 1529  
Qy 1308 GACACGGCCCTATACGTTTCATGGTGGCTTACAGGCTTTTACGTCACCAATGATACGGCT 1367  
Db 1530 GACACGGCCCTATACGTTTCATGGTGGCTTACAGGCTTTTACGTCACCAATGATACGGCT 1589  
Qy 1368 TGCAGATGATCTCTACCGATATCATGTGGATATCCAGATGTGACCATCTTAAAGACAG 1427  
Db 1590 TGCAGATGATCTCTACCGATATCATGTGGATATCCAGATGTGACCATCTTAAAGACAG 1649  
Qy 1428 CGGATTTTCCGTTACTTGCACACAGCTGTGATAGTGAAGTGGAAACCATCTGGTGTGG 1487  
Db 1650 CGGATTTTCCGTTACTTGCACACAGCTGTGATAGTGAAGTGGAAACCATCTGGTGTGG 1709  
Qy 1488 GCGAACAACACAAATGACACATCTATGACCATGCGGCGCAATGCTTCTCTCAGATTT 1547  
Db 1710 GCGAACAACACAAATGACACATCTATGACCATGCGGCGCAATGCTTCTCTCAGATTT 1769  
Qy 1548 CATGGCTATGACATTTGCCTGTGACCGCTGGTCAGTGTCTCCAGACCTGAT-TCACCA 1606  
Db 1770 CATGGCTATGACATTTGCCTGTGACCGCTGGTCAGTGTCTCCAGACCTGATCTCCACA 1829  
Qy 1607 TGATGTCAACAGATTTGGCCATTCACAGAGTCTTTACACAAAGCACCAATGATGTGG 1666  
Db 1830 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAAAGCACCAATGATGTGG 1888  
Qy 1667 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCTGGTATTTCACTCGGAACAGTGTGATG 1726  
Db 1889 GTGGTTTCAATAGTCTCCTCCTCAGCGACATCTGGTATTTCACTCGGAACAGTGTGATG 1948  
Qy 1727 CGCATCGAGTGAAGCGCTGTGTTTACGACGAGACCTGGTATTTGGTGTGTGGAACA 1786  
Db 1949 CGCATCGAGTGAAGCGCTGTGTTTACGACGAGACCTGGTATTTGGTGTGTGGAACA 2008  
Qy 1787 CAGGTCGTCTCAGTGTACTCTGGCGCTGGCACTCATGATGAAACAGGAAGAAAGTTAA 1846  
Db 2009 CAGGTCGTCTCAGTGTACTCTGGCGCTGGCACTCATGATGAAACAGGAAGAAAGTTAA 2068  
Qy 1847 AATCAGAAATTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACAGATT 1906  
Db 2069 AATCAGAAATTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACAGATT 2128  
Qy 1907 GTTACAGCTGCACAGCAACCAATGATGCGATGGTGCAATGATGATGTCCTCCCA 1966  
Db 2129 GTTACAGCTGTACAGCAACCAATGATGCGATGGTGCAATGATGATGTCCTCCCA 2188  
Qy 1967 GGAACACACAGCTCTCAGAAAGGCGAGATCTCAATTTTTAGTATGAGATTCGCCCAAGG 2026  
Db 2189 GGAACACACAGCTCTCAGAAAGGCGAGATCTCAATTTTTAGTATGAGATTCGCCCAAGG 2248  
Qy 2027 ATAAACCTTAGTACTCTGTAAACAGAACCAAGCTGACAGGAGCTGTGCCCTGGACCA 2086

Db 2249 ATAAACCTTAGTACTCTGTAAACAGAACCAAGCTGACAGGAGCTGTGCCCTGGACCA 2308  
Qy 2087 ACTGCCAGTGGAGCCCGGAATCAGGAGTGCAATGGCCCTCCCGCAAAATATCTGTGGCA 2146  
Db 2309 ACTGCCAGTGGAGCCCGGAATCAGGAGTGCAATGGCCCTCCCGCAAAATATCTGTGGCA 2368  
Qy 2147 TTGGCTGSCATTTGGTTGGAACTCATGTTTGAATAATTTACTCTGCCAAGGAGATTTATG 2206  
Db 2369 TTGGCTGSCATTTGGTTGGAACTCATGTTTGAATAATTTACTCTGCCAAGGAGATTTATG 2428  
Qy 2207 ACAATGCTAAATTTGTTGTAGGAACCAACAATGGCCCTTTTGGCTTCTCTTCAACCCAGA 2266  
Db 2429 ACAATGCTAAATTTGTTGTAGGAACCAACAATGGCCCTTTTGGCTTCTCTTCAACCCAGA 2488  
Qy 2267 AGAAGGTAGATTTGTTCTTAAGCAGCTGCGCAATAATGCAAGTCACTCAGAGCATGTCCA 2326  
Db 2489 AGAAGGTAGATTTGTTCTTAAGCAGCTGCGCAATAATGCAAGTCACTCAGAGCATGTCCA 2548  
Qy 2327 AGCTCACCTTAACCCCATGGGTCGGCTTCGGAAGATCAATGTGCTCTACTCTGGTGGG 2386  
Db 2549 AGCTCACCTTAACCCCATGGGTCGGCTTCGGAAGATCAATGTGCTCTACTCTGGTGGG 2608  
Qy 2387 AAGATATGTCCTCCATTTTACAAATAGTTTATACAGTGGATGCGCTGTGAGCCCATGTATG 2446  
Db 2609 AAGATATGTCCTCCATTTTACAAATAGTTTATACAGTGGATGCGCTGTGAGCCCATGTATG 2668  
Qy 2447 CTGGATTTCTGGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTCAACCTGCA 2506  
Db 2669 CTGGATTTCTGGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGCTCAACCTGCA 2728  
Qy 2507 TCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGTCAAAACCAAGTGTCTAAGCAGTGCC 2566  
Db 2729 TCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGTCAAAACCAAGTGTCTAAGCAGTGCC 2788  
Qy 2567 GGCACCATGTGCTTTGAGGACAGCATGTGAGATTCGACAGCGGAGCTCTCAGTGCA 2626  
Db 2789 GGCACCATGTGCTTTGAGGACAGCATGTGAGATTCGACAGCGGAGCTCTCAGTGCA 2848  
Qy 2627 TGTGTCGAGCAACATGAAGCAGTGTGTGACTCCAATGCTATGTGGCTCTCTTCCCTT 2686  
Db 2849 TGTGTCGAGCAACATGAAGCAGTGTGTGACTCCAATGCTATGTGGCTCTCTTCCCTT 2908  
Qy 2687 TTGGCAGTGTATGGAATGGTATACGATGAGCACTGCCCCCTTGAAAAATTTGTTCAGCT 2746  
Db 2909 TTGGCAGTGTATGGAATGGTATACGATGAGCACTGCCCCCTTGAAAAATTTGTTCAGCT 2968  
Qy 2747 ACTGTACTGTAGTCAATTGTGAGCAACAGGCTGTGGCTGTGTACTGTATCCACAGCA 2806  
Db 2969 ACTGTACTGTAGTCAATTGTGAGCAACAGGCTGTGGCTGTGTACTGTATCCACAGCA 3028  
Qy 2807 ATACTGCAAAAGGAAATGCAATAGAGGTTCTTATAAAGGACCAAGTGAAGATGCTTGGC 2866  
Db 3029 ATACTGCAAAAGGAAATGCAATAGAGGTTCTTATAAAGGACCAAGTGAAGATGCTTGGC 3088  
Qy 2867 AAGCCCTTACAGGAAATTTCTATCCACAGCCCTGTCTCAATCCAGCATGTGTCTAGAGG 2926  
Db 3089 AAGCCCTTACAGGAAATTTCTATCCACAGCCCTGTCTCAATCCAGCATGTGTCTAGAGG 3148  
Qy 2927 ACAGCAGATCAACTGCTTTTCAATTCACCTGTCCAGCTTGCATGCAACAGCCACAGTA 2986  
Db 3149 ACAGCAGATCAACTGCTTTTCAATTCACCTGTCCAGCTTGCATGCAACAGCCACAGTA 3208  
Qy 2987 AATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTTGACCAAGGAGCACTGGC 3046  
Db 3209 AATGATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACTTGACCAAGGAGCACTGGC 3268  
Qy 3047 AGACTGTCATATCTGGCTTTTACCGGTGATCCCAATGAGGGAATTTGTTCAGGCATGCA 3106  
Db 3269 AGACTGTCATATCTGGCTTTTACCGGTGATCCCAATGAGGGAATTTGTTCAGGCATGCA 3328  
Qy 3107 AGTGAATGGGCAAGCTCTCTGTGGAACCAACAGGGAAGTGTCTTGTGCAACCA 3166  
Db 3329 AGTGAATGGGCAAGCTCTCTGTGGAACCAACAGGGAAGTGTCTTGTGCAACCA 3388



This sequence represents a human mahogany gene of the invention. The mahogany genes are used: (i) to produce recombinant mahogany (mg) proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming therapeutics; (iii) as a source of diagnostic probes and primers for detecting expression of mg genes or mutations, regulatory defects, in this gene, or for isolation of related sequences; and (iv) in (cell-based) gene therapy. (II) are used to raise specific antibodies (Ab); to identify other (extra)cellular products involved in weight regulation, and to screen for agents that disrupt interaction between (II) and other macromolecules. The Ab are used to detect abnormal levels (or function) of (II) for diagnosis, prognosis or monitoring of treatment; to evaluate (II)-expressing cells intended for cell therapy, and as therapeutic mg inhibitors. Cells that express the mg gene (or contain the mg polypeptide) are used to identify agents (A) that modulate mg activity. (A) are potentially useful for the treatment of body weight disorders, particularly obesity, cachexia or anorexia, or other conditions associated with the mg gene such as hyperpigmentation, hyperphagia and disorders that result in increased metabolic rate

Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;

XX Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;

Query Match	96.7%;	Score 3476.8;	DB 3;	Length 8590;
Best Local Similarity	98.8%;	Pred. No. 0;		
Matches 3524; Conservative	0;	Mismatches 42;	Indels 2;	Gaps 2

Qy	15	GGCGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGCGGACGGCGAGCGCTCGCGGGCAG	74
Db	234	GGAGGCCGAGGCCGCTGCGGTGGCGGGCGGGGTGTCGGGCTCGGCGCGCAGCCGAGGCCAA	293
Qy	75	GAGCGCGGGCGCACTGCTGTCAACGGCGGTGCTGCGCAACCCCTGGCACCGGCCAGTGGGT	134
Db	294	GGAAATGTGACCGGCGGTGTGTCAAACGGCGGTGCTGCAACCCCTGGCACCGGCCAGTGGGT	353
Qy	135	CTGCCCGCGGCTGGGTGGCGAGCAATCCAGCACTGCGGGGCGGCTTCAGACTAAC	194
Db	354	CTGCCCGCGGCTGGGTGGCGAGCAATCCAGCACTGCGGGGCGGCTTCAGACTAAC	413
Qy	195	TGGATCTCTCGGTTTGTGACAGATGCACTGGAAATTATAATACAAACCGAGTGCAC	254
Db	414	TGGATCTCTCGGTTTGTGACAGATGCACTGGAAATTATAATACAAACCGAGTGCAC	473
Qy	255	GTGGCTCATTTGAAGACACGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGTCTAC	314
Db	474	GTGGCTCATTTGAAGACACGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGTCTAC	533
Qy	315	AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCAACCGCTAGT	374
Db	534	AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCAACCGCTAGT	593
Qy	375	TGCTGCATTTAGTGGCCCTCATGTTTCTGTGAGAGAGATGGCAATGAGACTGTCCCTCAGGT	434
Db	594	TGCTGCATTTAGTGGCCCTCATGTTTCTGTGAGAGAGATGGCAATGAGACTGTCCCTCAGGT	653
Qy	435	TGTTGCCACATCAGGTTATGCTGTGCTCATTTTTTTTAGTGAATGCTGCTTTATAATTGAC	494
Db	654	TGTTGCCACATCAGGTTATGCTGTGCTCATTTTTTTTAGTGAATGCTGCTTTATAATTGAC	713
Qy	495	TGGATTTAATATTACTACAGTTTGTATGTGTGCCAAATAACTGTCTCAGGCCGAGGAGA	554
Db	714	TGGATTTAATATTACTACAGTTTGTATGTGTGCCAAATAACTGTCTCAGGCCGAGGAGA	773
Qy	555	GTGTAGATTCAGTAATAGCAGCGATCTGTTTGAATGTGAATGTTTCTGAAACCTGGAAGG	614
Db	774	GTGTAGATTCAGTAATAGCAGCGATCTGTTTGAATGTGAATGTTTCTGAAACCTGGAAGG	833
Qy	615	TGAAGCATGTGCATTCCTCACTGTACAGACAACTGTGGTTTTTTCCTCATCAGGCGCATCTG	674
Db	834	TGAAGCATGTGCATTCCTCACTGTACAGACAACTGTGGTTTTTTCCTCATCAGGCGCATCTG	893
Qy	675	CAATTCAGTGTGTGACAGGATGCTCCTGCTTCTCAGACTGCGAGGTCCTGGATGTTTC	734
Db	894	CAATTCAGTGTGTGACAGGATGCTCCTGCTTCTCAGACTGCGAGGTCCTGGATGTTTC	953

3167	Qy	AGGCGCTCAAGGGGACGAGTGCACGACTATGTGAGGTAGAAAATCGATCAACAGGAACCC	3222
3389	Db	AGGCGCTCAAGGGGGACGAGTGCACGCTATGTGAGGTAGAAAATCGATCAACAGGAACCC	3448
3227	Qy	CTCTCAGAGAAACATGTTATTATTA	3286
3449	Db	CTCTCAGAGAAACATGTTATTATTA	3508
3287	Qy	CCGAGGAAGATGATCGCTATTATTA	3346
3509	Db	CCGAGGAAGATGATCGCTATTATTA	3568
3347	Qy	ACAGGATTTGGACATGCTTCATCAA	3406
3569	Db	ACAGGATTTGGACATGCTTCATCAA	3628
3407	Qy	CTGCCAGTTCCTCAGCTGGAAACCC	3466
3629	Db	CTGCCAGTTCCTCAGCTGGAAACCC	3688
3467	Qy	ACATTAGGAGTACAAAGATAGTTTCT	3526
3689	Db	ACATTAGGAGTACAAAGATAGTTTCT	3748
3527	Qy	ATATCACTTCTTTGTTTATGTTCAGT	3586
3749	Db	ATATCACTTCTTTGTTTATGTTCAGT	3808
3587	Qy	CTGAACCAATGA	3597
3809	Db	CTGAACCAATGA	3819

RESULT 7	
AAZ91920	
ID	AAZ91920 standard; cDNA; 8590 BP.
XX	
XX	
AC	AAZ91920;
XX	
XX	
DT	08-JUN-2000 (first entry)
XX	
DE	Human mahogany protein coding sequence #2.
XX	
XX	Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW	weight regulation; cell therapy; body weight disorder; cachexia;
KW	anorexia; hyperpigmentation; increased metabolic rate disorder;
XX	hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX	
OS	Homo sapiens.
XX	
PN	WO200005373-A2.
XX	
PD	03-FEB-2000.
XX	
XX	21-JUL-1999; 99WO-US016484.
XX	
PR	21-JUL-1998; 98US-0093630P.
PR	20-OCT-1998; 98US-0104978P.
PR	05-FEB-1999; 99US-00245041.
XX	
PA	(MILL-) MILLENIUM PHARM INC.
XX	
PI	Moore K, Nagle DL;
XX	
DR	WPI: 2000-195103/17.
DR	P-PSDB; AAY81807.
XX	
XX	New human and murine mahogany genes, useful, e.g. for diagnosis and
PT	treatment of body weight disorders.
PT	
XX	Claim 1; Fig 18a; 1989P; English.
PS	



QY 735 AGTTCCTGTACCACTAACAGTCAATTTTGGACTCGAGAGGAATATTCTTAACCTTAAGCT 794  
Db 954 AGTTCCTGTACCACTAACAGTCAATTTTGGACTCGAGAGGAATATTCTTAACCTTAAGCT 1013  
QY 795 CCCAGAGCACTCTATAAAGCTGTGGTCAATGTGAACCAATATATGTGGGTTGTGGAGGATA 854  
Db 1014 CCCAGAGCACTCTATAAAGCTGTGGTCAATGTGAACCAATATATGTGGGTTGTGGAGGATA 1073  
QY 855 TATGTTCAACCACTCACATATATACATGTTCTAGGATGATGACCTGCTCTTAGGAGTG 914  
Db 1074 TATGTTCAACCACTCACATATATACATGTTCTAGGATGATGACCTGCTCTTAGGAGTG 1133  
QY 915 GCTTCCACTAAACCGTTCTGTGAACCAATGTGGTGTGTAGATATGGTCAATCTTTGGCAAT 974  
Db 1134 GCTTCCACTAAACCGTTCTGTGAACCAATGTGGTGTGTAGATATGGTCAATCTTTGGCAAT 1193  
QY 975 ATACAGGATAAAATTTACATGTATGAGGAAAAATTTGATTCAACTGGGAAATGTGACCAA 1034  
Db 1194 ATACAGGATAAAATTTACATGTATGAGGAAAAATTTGATTCAACTGGGAAATGTGACCAA 1253  
QY 1035 TGAGTTGAGAGTATTTTACATTCATAATAGTCAATGGGTTGTTGACCCCTTAAGCAAA 1094  
Db 1254 TGAGTTGAGAGTATTTTACATTCATAATAGTCAATGGGTTGTTGACCCCTTAAGCAAA 1313  
QY 1095 GGAGCAGTATGAGTGGTGGGCACTCTGCACACATGTTTACACTGAAGAATGGCCGAGT 1154  
Db 1314 GGAGCAGTATGAGTGGTGGGCACTCTGCACACATGTTTACACTGAAGAATGGCCGAGT 1373  
QY 1155 GGTTCATCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATATAGCAATGTGCAGGA 1214  
Db 1374 GGTTCATCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATATAGCAATGTGCAGGA 1433  
QY 1215 ATATGATTTGGATAAGAACACATGGAGTATATACACACCAGGGTGCCCTTTGTGCAAGG 1274  
Db 1434 ATATGATTTGGATAAGAACACATGGAGTATATACACACCAGGGTGCCCTTTGTGCAAGG 1493  
QY 1275 GGTTACGGCCATAGCAGTGTGTTACGACCATAGGACAGGGCCCTATATAGTTTCATGGTGG 1334  
Db 1494 GGTTACGGCCATAGCAGTGTGTTACGACCATAGGACAGGGCCCTATATAGTTTCATGGTGG 1553  
QY 1335 CTACAAAGCTTTTCAGTGGCCAAATAGTACCGGCTTTGAGATGATCTCTACCGATATGATGT 1394  
Db 1554 CTACAAAGCTTTTCAGTGGCCAAATAGTACCGGCTTTGAGATGATCTCTACCGATATGATGT 1613  
QY 1395 GGATACCCAGATGTGACCAATCTTAAAGACAGCCCAATTTTCGGTTATCTTGACACACAGC 1454  
Db 1614 GGATACCCAGATGTGACCAATCTTAAAGACAGCCCAATTTTCGGTTATCTTGACACACAGC 1673  
QY 1455 TGTGATAGTGAAGTGAACCAATGCTGTGTTTGGGGGAAACACACACAATGACACATCTAT 1514  
Db 1674 TGTGATAGTGAAGTGAACCAATGCTGTGTTTGGGGGAAACACACACAATGACACATCTAT 1733  
QY 1515 GAGCCATGGCGCAAAATGCTTCTTACAGATTTTCATGGGCTATGACATTCGCTGTGACCG 1574  
Db 1734 GAGCCATGGCGCAAAATGCTTCTTACAGATTTTCATGGGCTATGACATTCGCTGTGACCG 1793  
QY 1575 CTGGTCACTGCTTCCAGACCTGAT-TCCACCAATGATGTCACAGATTTGGCCATTCAG 1633  
Db 1794 CTGGTCACTGCTTCCAGACCTGATCTCCACCAATGATGTCACAGATTTGGCCATTCAG 1852  
QY 1634 CAGTCTTACACAACAGCACCATATGATGTGCTGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693  
Db 1853 CAGTCTTACACAACAGCACCATATGATGTGCTGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912  
QY 1694 ACATCTGGTATTTCACTTCGAAACAGTGTGATGCGCATCGGATGGAAGCCGCTTGTATTAG 1753  
Db 1913 ACATCTGGTATTTCACTTCGAAACAGTGTGATGCGCATCGGATGGAAGCCGCTTGTATTAG 1972  
QY 1754 CAGCAGGACTGCTGATTTGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTGTGGG 1813  
Db 1973 CAGCAGGACTGCTGATTTGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTGTGGG 2032  
QY 1814 CGTGGCACTGTAGAACAAAGAAAAAGTTAAATCAGAAATGTTTTTCCAAAAAGAACTC 1873

Db 2033 CCGTGGCACTGATGAAACAGAGAAAAAGTTAAATCAGAAATGTTTTTCCAAAAAGAACTC 2092  
QY 1874 TTGACCATGACAGATGTGACCAAGCACACAGATTTGTTACAGCTGCAACCCAAACACCAATG 1933  
Db 2093 TTGACCATGACAGATGTGACCAAGCACACAGATTTGTTACAGCTGCAACCCAAACACCAATG 2152  
QY 1934 ACTGGCACTGTGTCAATGACCAATGTGTGCCAGGAACCAACAGCTGCTCAGAGGCCAGA 1993  
Db 2153 ACTGGCACTGTGTCAATGACCAATGTGTGCCAGGAACCAACAGCTGCTCAGAGGCCAGA 2212  
QY 1994 TCTCCATTTTATAGTATGAGAAATTCGCCCAAGGATAACCTATGTACTACTGTAAACAAGA 2053  
Db 2213 TCTCCATTTTATAGTATGAGAAATTCGCCCAAGGATAACCTATGTACTACTGTAAACAAGA 2272  
QY 2054 AGACAGCTGCAGGAGCTGTGCCCTGGACACAGAACTGCGAGTGGAGGCCCGGAATCAGG 2113  
Db 2273 AGACAGCTGCAGGAGCTGTGCCCTGGACACAGAACTGCGAGTGGAGGCCCGGAATCAGG 2332  
QY 2114 AGTGCATTTGCCCTGCCCGAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACTCAT 2173  
Db 2333 AGTGCATTTGCCCTGCCCGAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACTCAT 2392  
QY 2174 GTTTCAAAAATTTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACC 2233  
Db 2393 GTTTCAAAAATTTACTACTGCCAAGGAGAAATTTATGACAAATGCTAAATTTCTGTAGGAACC 2452  
QY 2234 ACAATGCCCTTTTGGCTTCTCTTCAACCCAGAGAGAGGTAGAAATTTGCTCTTAAGCAGC 2293  
Db 2453 ACAATGCCCTTTTGGCTTCTCTTCAACCCAGAGAGAGGTAGAAATTTGCTCTTAAGCAGC 2512  
QY 2294 TCGGAATTAATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGGCC 2353  
Db 2513 TCGGAATTAATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGGCC 2572  
QY 2354 TTCCGAAGATCAATGTGTCTTACTGCTGGGAGAGATATGTCCCCATTTTACAAATAGTT 2413  
Db 2573 TTCCGAAGATCAATGTGTCTTACTGCTGGGAGAGATATGTCCCCATTTTACAAATAGTT 2632  
QY 2414 TACTACAGTGGATGCGCTCTGAGGCCAGTGTGGAATTCCTGGAATTTTATCAGAAC 2473  
Db 2633 TACTACAGTGGATGCGCTCTGAGGCCAGTGTGGAATTCCTGGAATTTTATCAGAAC 2692  
QY 2474 CCAGTACTCGGGGAGTGAAGGCTGCACTGCACTCAACCCACTCAATGGTGTGTGTG 2533  
Db 2693 CCAGTACTCGGGGAGTGAAGGCTGCACTGCACTCAACCCACTCAATGGTGTGTGTG 2752  
QY 2534 AAAGGCTTGCAAAACACAGTGTGAAGCAGTCCCGACACCATGTGCTTGAAGGACAGCAT 2593  
Db 2753 AAAGGCTTGCAAAACACAGTGTGAAGCAGTCCCGACACCATGTGCTTGAAGGACAGCAT 2812  
QY 2594 GTGGAGATTGCAACAGGGGAGCTCTGAGTGCATGTGTGCAGCAACATGAAGCAGTGTG 2653  
Db 2813 GTGGAGATTGCAACAGGGGAGCTCTGAGTGCATGTGTGCAGCAACATGAAGCAGTGTG 2872  
QY 2654 TGGACTCCAAATGCTATGTGGGCTTCTTCCCTTTTGGCCAGTGTATGGAATGCTTATACGA 2713  
Db 2873 TGGACTCCAAATGCTATGTGGGCTTCTTCCCTTTTGGCCAGTGTATGGAATGCTTATACGA 2932  
QY 2714 TGAGCACCTGCCCCCTTGAAAAATTTGTTTCAGGCTACTGTACCTGTAGTCAATTTGTTGAGC 2773  
Db 2933 TGAGCACCTGCCCCCTTGAAAAATTTGTTTCAGGCTACTGTACCTGTAGTCAATTTGTTGAGC 2992  
QY 2774 AACCCGGTGTGGCTGTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCAATAGAGG 2833  
Db 2993 AACCCGGTGTGGCTGTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCAATAGAGG 3052  
QY 2834 GTTTCCTATAAAGGACCACTGGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCAC 2893  
Db 3053 GTTTCCTATAAAGGACCACTGGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCAC 3112  
QY 2894 AGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGCATACCACTGGTCTTTTCATTC 2953



Db 750 TCCAAATAACTGCTCAGGCCGAGGAGAGTGTAAAGTCACTAATAGCAGCGAAACTGTGCA 809  
Qy 588 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647  
Db 810 ATGTGAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869  
Qy 648 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCGAATTCAGTGAATGTCAGAGGATGCTCCTGCTT 707  
Db 870 CTGTGGTTTTCCTCATCGAGGCATCTGCAATTCGAATTCAGTGAATGTCAGAGGATGCTCCTGCTT 929  
Qy 708 CTCAGACTGCGAGGTCCTGAGATGTTCAAGTTCCTGTACAGCTTAACAGTCATTTTGGAC 767  
Db 930 CTCAGACTGCGAGGTCCTGAGATGTTCAAGTTCCTGTACAGCTTAACAGTCATTTTGGAC 989  
Qy 768 TCAGAGGAATAATCTTAACCTTAAAGCTCCCGAGCATCTCATAAAGCTGTGTCATAGG 827  
Db 990 TCAGAGGAATAATCTTAACCTTAAAGCTCCCGAGCATCTCATAAAGCTGTGTCATAGG 1049  
Qy 828 AAACATTTATGCGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 887  
Db 1050 AAACATTTATGCGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCT 1109  
Qy 888 AGCGTATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCGTTCGTGCAACAAATGGTGT 947  
Db 1110 AGCGTATGACCTTGTCTTAGGAGTGGCTTCCACTAAACCGTTCGTGCAACAAATGGTGT 1169  
Qy 948 TGTAGATATGTCATCTTTGGCAATATACAGGATTAATAATTTACATGTATGGAGAAA 1007  
Db 1170 TGTAGATATGTCATCTTTGGCAATATACAGGATTAATAATTTACATGTATGGAGAAA 1229  
Qy 1008 AATTGATTCAACTGGGAATGTGACCAATGAGTGTGAGAGTTCCTGCAACAAATGGTGT 1067  
Db 1230 AATTGATTCAACTGGGAATGTGACCAATGAGTGTGAGAGTTCCTGCAACAAATGGTGT 1289  
Qy 1068 ATGGGTGTTGTGACCCCTTAAGCAAAAGGAGGAGTATGAGTGGTGGCACTCTGCA 1127  
Db 1290 ATGGGTGTTGTGACCCCTTAAGCAAAAGGAGGAGTATGAGTGGTGGCACTCTGCA 1349  
Qy 1128 CATTTGTACACTGAAGTGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187  
Db 1350 CATTTGTACACTGAAGTGGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1409  
Qy 1188 CTATGATATATAAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATAT 1247  
Db 1410 CTATGATATATAAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATAT 1469  
Qy 1248 ACACACCGAGGTGCTTGTGCAAGGGGTTACGCCCATAGCAGTGTTCAGCACTAG 1307  
Db 1470 ACACACCGAGGTGCTTGTGCAAGGGGTTACGCCCATAGCAGTGTTCAGCACTAG 1529  
Qy 1308 GACCAGGGCCCTATACGTTTCATGCTGCTACAGGCTTTTCAGTGCCCAATAGTACCGCT 1367  
Db 1530 GACCAGGGCCCTATACGTTTCATGCTGCTACAGGCTTTTCAGTGCCCAATAGTACCGCT 1589  
Qy 1368 TGCAATGATCTCTACCGATATGATGGATATGCAATGATGATGATGATGATGATGATGAT 1427  
Db 1590 TGCAATGATCTCTACCGATATGATGGATATGCAATGATGATGATGATGATGATGATGAT 1649  
Qy 1428 CCGATTTTCCGTTACTTGCACACAGCTGTGATGATGATGATGATGATGATGATGATGAT 1487  
Db 1650 CCGATTTTCCGTTACTTGCACACAGCTGTGATGATGATGATGATGATGATGATGATGAT 1709  
Qy 1488 GGGAAACACACAATGACACATCTATGAGCCATGCGCCAAATGCTTCTCTTCAGATTT 1547  
Db 1710 GGGAAACACACAATGACACATCTATGAGCCATGCGCCAAATGCTTCTCTTCAGATTT 1769  
Qy 1548 CATGGCTATGACATTCCTGTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1606  
Db 1770 CATGGCTATGACATTCCTGTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1829  
Qy 1607 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAAACAGCACCATGTATGTTGCG 1666  
Db 1830 TGATGTCAACAGATTTGGCCATTCCAGCAGTCTTACACAAACAGCACCATGTATGTTGCG 1888

Qy 1657 GTGGTTTCAATAGTCTCTCTCCTCAGCGACATCTCTGGTATTCACCTCGGAAACAGTGTGATG 1726  
Db 1889 GTGGTTTCAATAGTCTCTCTCCTCAGCGACATCTCTGGTATTCACCTCGGAAACAGTGTGATG 1948  
Qy 1727 CGCATCGAGTGAAGCGCTTGTGTAGCAGGACCTGCTGATTCGGTGTGTGTGGAAACA 1786  
Db 1949 CGCATCGAGTGAAGCGCTTGTGTAGCAGGACCTGCTGATTCGGTGTGTGTGGAAACA 2008  
Qy 1787 CAGGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAAAGTTAA 1846  
Db 2009 CAGGTCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAAAGTTAA 2068  
Qy 1847 AATCAGATGTTTTCAAAAAGAACTCTTGACCAATGACAGATGTGACAGCACAGATT 1906  
Db 2069 AATCAGATGTTTTCAAAAAGAACTCTTGACCAATGACAGATGTGACAGCACAGATT 2128  
Qy 1907 GTTACAGCTGACAGCAACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1966  
Db 2129 GTTACAGCTGACAGCAACCAATGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2188  
Qy 1967 GGAACACAGCTGCTCAGAGGCCAGATCTCCATTTTGTAGTATGAGAAATGCCCCAAGG 2026  
Db 2189 GGAACACAGCTGCTCAGAGGCCAGATCTCCATTTTGTAGTATGAGAAATGCCCCAAGG 2248  
Qy 2027 ATACCTCTATGCTACTGTHAACAGAGACCACTGAGAGCTGTGAGAGCTGTGCTGAGACAGA 2086  
Db 2249 ATACCTCTATGCTACTGTHAACAGAGACCACTGAGAGCTGTGAGAGCTGTGCTGAGACAGA 2308  
Qy 2087 ACTGCACTGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCCCGGAAATATCTGTGGCA 2146  
Db 2309 ACTGCCAGTGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCCCGGAAATATCTGTGGCA 2368  
Qy 2147 TTGGCTGCACTTTGGTGGAACTCATGTTTGGAAATTAATCTGCTGCAAGAGAAATATG 2206  
Db 2369 TTGGCTGCACTTTGGTGGAACTCATGTTTGGAAATTAATCTGCTGCAAGAGAAATATG 2428  
Qy 2207 ACAATGCTAAATGCTCTGTAGGAACCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2266  
Db 2429 ACAATGCTAAATGCTCTGTAGGAACCAACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2488  
Qy 2267 AGAAGGTAGAAATTTGCTCTTAAGCAGCTGCGAAATTAATGCACTCATCTCAGAGCATGTCCA 2326  
Db 2489 AGAAGGTAGAAATTTGCTCTTAAGCAGCTGCGAAATTAATGCACTCATCTCAGAGCATGTCCA 2548  
Qy 2327 AGCTCACTTAACCCCATGGTGGCTGCGCTGCGAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2386  
Db 2549 AGCTCACTTAACCCCATGGTGGCTGCGCTGCGAGAGCAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2608  
Qy 2387 AAGATATGCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGTATG 2446  
Db 2609 AAGATATGCTCCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGTATG 2668  
Qy 2447 CTGATTTCTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAAGCTGCAACCTGTCCA 2506  
Db 2669 CTGATTTCTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAAGCTGCAACCTGTCCA 2728  
Qy 2507 TCAACCCACTCAATGTTAGTGTCTGTGAAAGGCTGCAAAACCAAGTGTGCAAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2566  
Db 2729 TCAACCCACTCAATGTTAGTGTCTGTGAAAGGCTGCAAAACCAAGTGTGCAAGCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2788  
Qy 2567 GGACACCATGCTGCTTGGAGCAGCATGTGGAGATTGCAACAGCGGAGCTCTGAGTGTCA 2626  
Db 2789 GGACACCATGCTGCTTGGAGCAGCATGTGGAGATTGCAACAGCGGAGCTCTGAGTGTCA 2848  
Qy 2827 TGTGGTGCAGCAACATGAGCAGTGTGGGACTTCCAATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2886  
Db 2849 TGTGGTGCAGCAACATGAGCAGTGTGGGACTTCCAATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2908  
Qy 2887 TTGGCCAGTGTATGGAATGTTATACAGTGTGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2746  
Db 2909 TTGGCCAGTGTATGGAATGTTATACAGTGTGAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2968

```
QY 2747 ACTGTACTCTAGTCAATTCCTGGAGCAACAGCGCTGTGGCTGTACTGTATCCAGCA 2806
Db 2969 ACTGTACTCTAGTCAATTCCTGGAGCAACAGCGCTGTGGCTGTACTGTATCCAGCA 3028
QY 2807 ATACTGCAAGGAAATCATAGAGGTTCTTATAAGGACAGTGAAGATGCTTTCGC 2866
Db 3029 ATACTGCAAGGAAATCATAGAGGTTCTTATAAGGACAGTGAAGATGCTTTCGC 3088
QY 2867 AAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTCACAGCATGTGTCTAGAG 2926
Db 3089 AAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTCACAGCATGTGTCTAGAG 3148
QY 2927 ACAGCAGATCAACTGTCTTTTCAATTCAGTCCAGTTCGCCAATCAACGGCCACAGTA 2986
Db 3149 ACAGCAGATCAACTGTCTTTTCAATTCAGTTCGCCAATCAACGGCCACAGTA 3208
QY 2987 AATGCATCAATCAGAGCATCTGTAGAGTGTGAGAACCTGACACAGGACAGCACTGCG 3046
Db 3209 AATGCATCAATCAGAGCATCTGTAGAGTGTGAGAACCTGACACAGGACAGCACTGCG 3268
QY 3047 AGACCTCATATCTGGCTTCTACGGTGATCCCAACCAATGAGGAAATGTACGCCATGCA 3106
Db 3269 AGACCTCATATCTGGCTTCTACGGTGATCCCAACCAATGAGGAAATGTACGCCATGCA 3328
QY 3107 AGTGCANTGGGACGCTCTCTGTGMAACCAACAGCGGCAAGTGTCTTGACACCA 3166
Db 3329 AGTGCANTGGGACGCTCTCTGTGMAACCAACAGCGGCAAGTGTCTTGACACCA 3388
QY 3167 AGGCGTCAAGGGGAGGAGTGCAGCTATGTGAGGTAGAAATTCATACCAAGGAAC 3226
Db 3389 AGGCGTCAAGGGGAGGAGTGCAGCTATGTGAGGTAGAAATTCATACCAAGGAAC 3448
QY 3227 CTCTCAGAGGAACATGTTATTAATCTTCTTATTCAGTATCATCATCAGTTTACCTTAT 3286
Db 3449 CTCTCAGAGGAACATGTTATTAATCTTCTTATTCAGTATCATCATCAGTTTACCTTAT 3508
QY 3287 CCAGGAGATGATCGCTATTACAGCTATCAATTTGTGGCTACTCTCGACGACAA 3346
Db 3509 CCAGGAGATGATCGCTATTACAGCTATCAATTTGTGGCTACTCTCGACGACAA 3568
QY 3347 ACAGGGATTTGGACATGTTTCATCAATGCTCCCAAGAAATTTCAACCTCAACATCAGCTGG 3406
Db 3569 ACAGGGATTTGGACATGTTTCATCAATGCTCCCAAGAAATTTCAACCTCAACATCAGCTGG 3628
QY 3407 CTGCGACTTCTCAGCTGGAAACCCAGCTCGAGAGAGATGCTGTGTTTCAAAAACCA 3466
Db 3629 CTGCGACTTCTCAGCTGGAAACCCAGCTCGAGAGAGATGCTGTGTTTCAAAAACCA 3688
QY 3467 ACATTAAAGGAGTACAAAGATAGTTTCTTAATGAGAAAGTTTGATTTTGGCAACCCAA 3526
Db 3689 ACATTAAAGGAGTACAAAGATAGTTTCTTAATGAGAAAGTTTGATTTTGGCAACCCAA 3748
QY 3527 ATATCACTTTCTTTTGTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGGT 3581
Db 3749 ATATCACTTTCTTTTGTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGGT 3803
```

## RESULT 9

ID AAS72659 standard; cdna; 8561 BP.

AA 8561 BP.

AC 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

XX 8561 BP.

```
XX 11-OCT-2001.
PD 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
PI WPI; 2001-639362/73.
XX P-PSDB; ABG08472.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
```

Claim 1; SEQ ID NO 8463; 103pp; English.

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations and to produce other types of data and products dependent on DNA and amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic coding sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

Query Match 96.5%; Score 3469.4; DB 5; Length 8561;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 3507; Conservative 0; Mismatches 26; Indels 2; Gaps 2;

```
QY 48 GCGCGGACGCGCAGCGCTCGCGGCGAGGCGCGGCGCGCAGCTGTGTCAACGCGGTGC 107
Db 270 GTGCGGCTCAGCGCGAGCGCGGCGCGGCGCGGCGCGCTGTGTCAACGCGGTGC 329
QY 108 CTGCACACCTGCGCAGCGCGCAGTGCCTGTGCCCGCGCGCTGGGTGGCGGAGCAATGCCA 167
Db 330 CTGCAACCTGCGCAGCGCGCAGTGCCTGTGCCCGCGCGCTGGGTGGCGGAGCAATGCCA 389
QY 168 GCACCTGCGGCGCGCGCTTCACAGCTAACTGGATCTCTGGGTTTGTGACAGATGACCTGG 227
Db 390 GCACCTGCGGCGCGCGCTTCACAGCTAACTGGATCTCTGGGTTTGTGACAGATGACCTGG 449
QY 228 AAATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 287
Db 450 AAATTATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 509
QY 288 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGGACCATTTATATGTTATGA 347
Db 510 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGGACCATTTATATGTTATGA 569
QY 348 TGGGAGCTCAATTTATGACACCGCTAGTGTGTCATTTAGTGGCCCTCTTGTTCCTGAGAG 407
Db 570 TGGGAGCTCAATTTATGACACCGCTAGTGTGTCATTTAGTGGCCCTCTTGTTCCTGAGAG 629
```

Qy	408	AGATGGCAATGAGACTGCTCCCTGAGGTTGTTGCCACATCAGGTTATGCCCTTGCTGCATTT	467
Db	630	AGATGGCAATGAGACTGCTCCCTGAGGTTGTTGCCACATCAGGTTATGCCCTTGCTGCATTT	689
Qy	468	TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAAATATTACTTTACAGTTTTCATATGTG	527
Db	690	TTTTAGTGATGCTGCTTATAATTTGACTGGATTTAAATATTACTTTACAGTTTTCATATGTG	749
Qy	528	TCCAAATAACTGCTCAGGCCGAGAGAGAGTGAAGATCAGTAATAATGACAGCATCTGTTGA	587
Db	750	TCCAAATAACTGCTCAGGCCGAGAGAGTGAAGATCAGTAATAATGACAGCATCTGTTGA	809
Qy	588	ATGTGAATGTCTCTGAAAACCTGAAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA	647
Db	810	ATGTGAATGTCTCTGAAAACCTGAAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA	869
Qy	648	CTGTGGTTTTCCTCATCTGAGGCATCTGCAATTTCAAGTGTATGATGTCTGAGGATGCTCTGCTT	707
Db	870	CTGTGGTTTTCCTCATCTGAGGCATCTGCAATTTCAAGTGTATGATGTCTGAGGATGCTCTGCTT	929
Qy	708	CTCAGACTGGCAGGGTCCTGGATGTTTCAGTTTCTCTGTACCAAGCTAACCAAGTCATTTTGGAC	767
Db	930	CTCAGACTGGCAGGGTCCTGGATGTTTCAGTTTCTCTGTACCAAGCTAACCAAGTCATTTTGGAC	989
Qy	768	TCGAGAGGATATTTCTAACTTAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGG	827
Db	990	TCGAGAGGATATTTCTAACTTAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGG	1049
Qy	828	AAACATTTATGTGGTGTGTGGAGGATATATGTTTCAACCACTTCAGATTATTAACATGTTCT	887
Db	1050	AAACATTTATGTGGTGTGTGGAGGATATATGTTTCAACCACTTCAGATTATTAACATGTTCT	1109
Qy	888	AGCCTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAAACGGTCTCTGTGAACCAATGTGT	947
Db	1110	AGCCTATGACCTTGCTTCTAGGGAGTGGCTTCCACTAAAACGGTCTCTGTGAACCAATGTGT	1169
Qy	948	TGTTAGATATGGTCAATCTTTTGGCAATATACAAGGATAAAATTTACATGTATCGAGGAAA	1007
Db	1170	TGTTAGATATGGTCAATCTTTTGGCAATATACAAGGATAAAATTTACATGTATCGAGGAAA	1229
Qy	1008	AATTGATTCACCTGGGAATGTGACCAATGAGTTGAGAGTTTTTTCATATTCATATGAGTC	1067
Db	1230	AATTGATTCACCTGGGAATGTGACCAATGAGTTGAGAGTTTTTTCATATTCATATGAGTC	1289
Qy	1068	ATGGGTGTGTGTGACCTTAGGCAAGGAGGAGTATGCAAGTGGTGGGACATCTGCACA	1127
Db	1290	ATGGGTGTGTGTGACCTTAGGCAAGGAGGAGTATGCAAGTGGTGGGACATCTGCACA	1349
Qy	1128	CATTGTTTACACTGAAGAAATGCGCGAGTGGTCATGCTGGTCATCTTTTGGTCACCTGCCCTCT	1187
Db	1350	CATTGTTTACACTGAAGAAATGCGCGAGTGGTCATGCTGGTCATCTTTTGGTCACCTGCCCTCT	1409
Qy	1188	CTATGGATATATAGCAATGTGCAGGAATATGATTTGGATAAGNACATCGGAGTATATT	1247
Db	1410	CTATGGATATATAGCAATGTGCAGGAATATGATTTGGATAAGNACATCGGAGTATATT	1469
Qy	1248	ACACACCCAGGGTCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTTACCAACCATAG	1307
Db	1470	ACACACCCAGGGTCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTTACCAACCATAG	1529
Qy	1308	GACCAGGGCCCTATACGTTTCAATGTGTGCTTACAGGCTTTTCAGTGCCCAATTAAGTACGGCT	1367
Db	1530	GACCAGGGCCCTATACGTTTCAATGTGTGCTTACAGGCTTTTCAGTGCCCAATTAAGTACGGCT	1589
Qy	1368	TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGACCAATCTTAAAGACAG	1427
Db	1590	TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGACCAATCTTAAAGACAG	1649
Qy	1428	CCGATTTTTTCGTTACTTTGACACAGCTGTGATAGTGAGTGGAAACCATGCTGGTGTGG	1487
Db	1650	CCGATTTTTTCGTTACTTTGACACAGCTGTGATAGTGAGTGGAAACCATGCTGGTGTGG	1709

QY	1488	GGGNAACACACACAAATGACACATCTATGAGCATGTGGCGCCAAATGCTTCTCTTCAGATTT	1547
Db	1710	GGGAAAACACACAAATGACACATCTATGAGCATGTGGCGCCAAATGCTTCTCTTCAGATTT	1769
QY	1548	CATGGCCCTATGACATTCGCTGTGACCCGCTGGTCAGTGCCTTCCACAGACCTGAT--TCCACCA	1606
Db	1770	CATGGCCCTATGACATTCGCTGTGACCCGCTGGTCAGTGCCTTCCACAGACCTGATCTCCACCA	1829
QY	1607	TGATGTCAAACAGATTTGGCCATTTCCAGCAGTCTTTACACAAACAGCACCAATGATGTGTTGCG	1666
Db	1830	TGATGTCAAACAGATTTGGCCATTT--CAGCAGTCTTTACACAAACAGCACCAATGATGTGTTGCG	1888
QY	1667	GTGGTTTCAATAGTCTCCTCCTCAGCGACATCTCGTGTATTCACTCGGAACAGTGTGATG	1726
Db	1889	GTGGTTTCAATAGTCTCCTCCTCAGCGACATCTCGTGTATTCACTCGGAACAGTGTGATG	1948
QY	1727	CGCATCGGAGTGAAGCCGCTGTTTACACAGAGAACCTGCTATTCGGTGTGTGTGGAACA	1786
Db	1949	CGCATCGGAGTGAAGCCGCTGTTT--TAGCAGCAGAACCTGCTATTCGGTGTGTGTGGAACA	2008
QY	1787	CAGGGTCGTCTCAGTGTATCTCGTGGCGCTGACCACTGATGAACAAAGAAAGAAAGTTAA	1846
Db	2009	CAGGGTCGTCTCAGTGTATCTCGTGGCGCTGACCACTGATGAACAAAGAAAGAAAGTTAA	2068
QY	1847	AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATTGACAGATGTGACCAGCACACAGATT	1906
Db	2069	AATCAGAATGTTTTTCCAAAAGAACTCTTGACCATTGACAGATGTGACCAGCACACAGATT	2128
QY	1907	GTTTACAGCTGCACAGCCAAACACCAATGACTGCCACTGGTGCATGACCATTTGTGCCCA	1966
Db	2129	GTTTACAGCTGTACAGCCAAACACCAATGACTGCCACTGGTGCATGACCATTTGTGCCCA	2188
QY	1967	GGAAACACAGCTGCTCAGAAAGCCAGATCTCCATTTTTAGGTATGAAATTGCCCCAAGG	2026
Db	2189	GGAAACACAGCTGCTCAGAAAGCCAGATCTCCATTTTTAGGTATGAAATTGCCCCAAGG	2248
QY	2027	ATAACCCCTATGTACTACTGTAAACAGAAACAGAGTCGACGAGGTGHCCTTGGAACGAGA	2086
Db	2249	ATAACCCCATGTACTACTGTAAACAGAAACAGAGTCGACGAGGTGHCCTTGGAACGAGA	2308
QY	2087	ACTGCCAGTGGAGCCCCCGGAATCAGAGGTGCATTGCCCTGCCGGAATAATCTGTGGCA	2146
Db	2309	ACTGCCAGTGGAGCCCCCGGAATCAGAGGTGCATTGCCCTGCCGGAATAATCTGTGGCA	2368
QY	2147	TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAATAATTACTATGCCAAGGAAATTTATG	2206
Db	2369	TTGGCTGGCATTTGGTTGGAAACTCATGTTTGAATAATTACTATGCCAAGGAAATTTATG	2428
QY	2207	ACAATGCTAAATTTGTTGTAGGAACACAAATGCCCTTTTGGCTTCTCTTACAAACCCAGA	2266
Db	2429	ACAATGCTAAATTTGTTGTAGGAACACAAATGCCCTTTTGGCTTCTCTTACAAACCCAGA	2488
QY	2267	AGAAGGTAGAAATTTGTCTTTAAGCAGCTGCGAATAATGCAGTCAICTCAGAGCATGTCCA	2326
Db	2489	AGAAGGTAGAAATTTGTCTTTAAGCAGCTGCGAATAATGCAGTCAICTCAGAGCATGTCCA	2548
QY	2327	AGTCACTTTAAACCCCATGGGTGGCCCTTCGGAAGATCAATGTCCTACTGTGTGCTGGG	2386
Db	2549	AGTCACTTTAAACCCCATGGGTGGCCCTTCGGAAGATCAATGTCCTACTGTGTGCTGGG	2608
QY	2387	AAGATATGTCCTCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCCAGTGA	2446
Db	2609	AAGATATGTCCTCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCCAGTGA	2668
QY	2447	CTGGATTCCTGGAAATTTTATACAAACCCAGTACTCGGGACTGGAAGGCTGCAACCTGCA	2506
Db	2669	CTGGATTCCTGGAAATTTTATACAAACCCAGTACTCGGGACTGGAAGGCTGCAACCTGCA	2728
QY	2507	TCAACCCACTCAATGCTAGTGTCTGTGAAGGCGCTGCACACACAGTGTCTAAGCATGTC	2566
Db	2729	TCAACCCACTCAATGCTAGTGTCTGTGAAGGCGCTGCACACACAGTGTCTAAGCATGTC	2788
QY	2567	GGACACCATCTGCCCTTTGAGACACAGCATGTGGAGATTGCACACAGCGGCAGCTCTGAGTGA	2626





Db 306 GGAGGCGGAGCGCTGCGGTGGCGGGCGTGTGCAGGCTGGCGCGCAGCTGAGGCCAA 365  
Qy 75 GAGCGGCGGCGCACTGTGTCAACGGCGGTGCTGCAACCCCTGGCAGCCGCGAGTGCCT 134  
Db 366 GGAATGTGACCGCGCGTGTCAACGGCGCGCTGCACACCCCTGGCAGCCGCGAGTGCCT 425  
Qy 135 CTCGCCCGCGCGTGGTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAAC 194  
Db 426 CTCGCCCGCGCGTGGTGGCGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAAC 485  
Qy 195 TGGATCTTCTGGGTTGTGACAGATGGACCTGGAAATATAAATACAAAACGAAGTGCAC 254  
Db 486 TGGATCTTCTGGATTGTAAACAGATGGACCTGGAAATATAAATACAAAACGAATGCAC 545  
Qy 255 GTGGCTCATTTGAAGGACAGCAAAATAGATAGACATTCGTTTCAATCATTTGCTAC 314  
Db 546 ATGGCTCATTTGAAGGACAGCAAAATAGATAGACATTCGATTCAACCAATTTGCTAC 605  
Qy 315 AGAGTGTAGTGGGACCATTTATATCTTTATGATGGGACTCAATTTATGCACCGTAGT 374  
Db 606 TGAATGTAGTGGACCATTTATATCTTTATGATGGGACTCAATTTATGCACCTCTGAT 665  
Qy 375 TGTGCTCATTTAGTGGCTCATTTCTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434  
Db 666 TGTGCTCATTTAGTGGCTCATTTCTCTGAGAGAGATGGCAATGAGAGCGCCCTGAGGT 725  
Qy 435 TGTGCGACATCAGTATGCTTGTGCTGCTATTTTATGATGGTCTTATATTTGAC 494  
Db 726 CACTGTCACTTCAAGTATGCACTACTGCTATTTTTCAGTGTGCTTATATTTGAC 785  
Qy 495 TGGATTATATTTACATTTACAGTTTTCATATGTCTCCAAATAACTGTCAAGCGCGAGGAGA 554  
Db 786 TGGATTATATCACATTACAAATTTGACATGTGCCGGAATATGCTCAGGCCGAGGAGA 845  
Qy 555 GTGTAAAGATCAGTAATAGCAGGATACCTGTTGAAATGTGAAATGTTCTGAAATGCGAAAGG 614  
Db 846 ATGTAAAGACAGTAAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAATGCGAAAGG 905  
Qy 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGTTTTCCTCATCGAGGATCTG 674  
Db 906 GAGTCTGTGACATTCCTCACTGTACAGACAACTGTGTTTTCCTCATCGAGGATCTG 965  
Qy 675 CAATTCAAGTGTGTGAGGATGCTCTCTCTCAGACTGGCAGGGTCTGGATGTTTC 734  
Db 966 TAAACGAACTGACACAGAGGGTGTCTCTCTCTCTCACTGGCAGGGTCTGGATGTTTC 1025  
Qy 735 AGTTCCTGTACAGCTAAACAGTCACTTTTGGACTCGAGGAAATATTCTAACTTAAAGCT 794  
Db 1026 AATTCTGTGCCAGCTTAAACAGTCTTTTGGACTCGAGGAAATATTCTGACTTAAAGCT 1085  
Qy 795 CCCAGAGCATCTCATAAAGCTGTGGTCAATGGAACAATATGTGGGTTGTTGGAGGATA 854  
Db 1086 TCCACAGAGCTCTCATAAAGCTGTGGTCAATGGAATAATATGTGGGTTGTTGGCGGTA 1145  
Qy 855 TATGTTCAACCACTCAGATTAACATGTTCTAGCGTATGACCTTGTCTTAGGGAGTG 914  
Db 1146 TATGTTCAACCACTCAGATTAACAGTATGTTTATAGCGTATGACCTGGCTTCTAGGGAATG 1205  
Qy 915 GCTTCCACTAAACCGTTCTGTGAACAATGTGTGTTGTAGATATGTTCTTTCTGGCAAT 974  
Db 1206 GCTTTCATTAACCATCTGTGNAACAGTGTGTGTGTGAATATGTTCTTTCTGGCAAT 1265  
Qy 975 ATACAGGATAAATTTATGATATGAGGAAATTTGATTTCAACTGGGAATGTACCAA 1034  
Db 1266 ACATAAGGATAAATCTACATGTATGAGGAAATTTGATTTCAACAGGGAATGTACCAA 1325  
Qy 1035 TGAAGTTGAGGTTTTTCACTTCAATAATGAGTATGGGTGTTGTGACCCCTTAAGCCAA 1094  
Db 1326 CGAGCTGAGAGTATTTCACTATTCAATAATGAATCATGGGTATTTGTTAAACCCCAAGCTAA 1385  
Qy 1095 GGAGCAGTATGAGTGGTGGGCACTCTGCACACATTTGTACACTGAAGAAATGGCGAGT 1154  
Db 1386 GGATCAGTATGAGTGGTGGGCACTCTGCACACATTTGTACACTGTGATCTGCGCGGTGT 1445

Qy 1155 GGTATGCTGCTCATCTTTGGTCACTGCCCCTCTCTATGGATATATAAGCAATGTGCAGGA 1214  
Db 1446 GGTATGCTGCTCATCTTCGGTCACTTGTCCGCTCTACGGATATATAAGTGTGTGCAGGA 1505  
Qy 1215 ATATGATTTGATTAAGAACACATGGAGTATATTACACCCAGGTCCTTTGTGCAGG 1274  
Db 1506 ATATGATTTGATTAAGAACACATGGAGTATATTACAACTCAAGTGTCTCTAGTGCAGG 1565  
Qy 1275 GGGTACGGCCATAGCAGTGTTTACGACCATAGGACAGGGCCCTATACGTTTCATGGTG 1334  
Db 1566 GGGTATAGCCACACAGAGTGTTTATGATCAGAGCAAGGCTCTGTATGTTTCATGGTG 1625  
Qy 1335 CTACAGGCTTTAGTGCATAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394  
Db 1626 CTACAGGCTTTAGTGCATAATAGTACCGGCTTGCAGATGATCTCTACAGATACCATGT 1685  
Qy 1395 GGATACCCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTTCGGTTACTTGCACACAGC 1454  
Db 1686 GGACACTCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTTCGGTTACTTGCACACAGC 1745  
Qy 1455 TGTGATGATGAGTGGAAACCATGTGTTTGGGGGAAACACACAAATGACACATCTAT 1514  
Db 1746 TGTGATGATGAGTGGAAACCATGTGTTTGGGGGAAACACACAAATGACACATCTAT 1805  
Qy 1515 GAGCATGGCGCCAAATGCTTCTCTCAGATTTTCATGGCCTATGACATTCGCTGTGACCG 1574  
Db 1806 GAGCATGGCGCCAAATGCTTCTCTCAGATTTTCATGGCCTATGACATTCGCTGTGACCG 1865  
Qy 1575 CTGCTCAGTGTCTCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCATTCAG 1633  
Db 1866 ATGCTCAGTGTCTCCAGACCTGAGCTCCATCAGATGTCAACCGATTTGGCCACT-CAG 1924  
Qy 1634 CAGTCTTACAAACAGACACATGATGTGTTGGTGTGTTTCAATGATGCTCTCTCAGCG 1693  
Db 1925 CAGTCTTACAAACAGACACATGATGTGTTGGTGTGTTTCAACAGCTCTCTCAGTG 1984  
Qy 1694 ACATCTGGTATTACACCTCGSAAACAGTGTGATGCGCATCGAGTGAAGCGCTGTTTAG 1753  
Db 1985 ACCTCTTAGTCTTCACTCGSAGCAGTGTGATGACACCGCAGCAGAGTGTCTTGTGG 2044  
Qy 1754 CAGCAGACCTGTGATTTGGTGTGTGGAAACAGGGTCTCTCAGTGTATCTGTTGGG 1813  
Db 2045 CAGCAGACCTGTGATTCGGTGTCTGTGGGACACACAGTGTCTCGATGTACCTCTCTGG 2104  
Qy 1814 CGCTGGCAACTGTATGAAAGAAAGTAAATCAGAAATGTTTTCCTTAAAGAACTC 1873  
Db 2105 AGTTGGCAACTGTATGAAAGAAAGTAAATCAGAAATGTTTTCCTTAAAGAACTC 2164  
Qy 1874 TTGACCATGACAGATGTGACAGCACAGATTTGTTACAGCTGCACAGCCAAACCAATG 1933  
Db 2165 TTGACCATGACAGATGTGACAGCACAGATTTGTTACAGCTGCACAGCCAAATACCAATG 2224  
Qy 1934 ACTGCACTGGTGAATGACATTTGTGTCAGGAAACACAGCTGTCTGAGAGCCAGGA 1993  
Db 2225 ACTGCACTGGTGAATGACATTTGTGTCAGTGTGTCAGTGTGTCAGAGAGCCAGGA 2284  
Qy 1994 TCTCCATTTTAGTATGAGATTTGCCCAAGGATTAACCTCTACTCTACTGTAAACAAGA 2053  
Db 2285 TCTCCATTTTAGTATGAGATTTGCCCAAGGATTAACCTCTACTCTACTGTAAACAAGA 2344  
Qy 2054 AGACAGCTGAGAGCTGTGCTGACAGCAAGTCTGCAAGTGGGAGCCCGGAAATCAGG 2113  
Db 2345 AAACAGCTGAGAGCTGTGCTGACAGCAAGTCTGCAAGTGGGAGCCCTCGAAATCAAG 2404  
Qy 2114 AGTGCATTTCCCTCGGAAATATCTGTGCAATGGCTGCAATTTGTTGGTGGAACTCAT 2173  
Db 2405 AGTGCATTTCCCTCGGAAATATCTGTGCAATGGCTGCAATTTGTTGGTGGAACTCAT 2464  
Qy 2174 GTTTGAAATTTACTCTGCAAGGAGAAATATGACAAATGCTTAAATTTCTGTAGGAACC 2233  
Db 2465 GTCTGAAATTTACTCTGCAAGGAGAAATATGACAAATGCTTAAATTTCTGTAGGAACC 2524





CC activity. (A) are potentially useful for the treatment of body weight  
CC disorders, particularly obesity, cachexia or anorexia, or other  
CC conditions associated with the mg gene such as hyperpigmentation,  
CC hyperphagia and disorders that result in increased metabolic rate  
XX  
SQ Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;

Query Match 79.1%; Score 2843.6; DB 3; Length 8827;  
Best Local Similarity 90.2%; Pred. No. 0;  
Matches 3063; Conservative 0; Mismatches 329; Indels 2; Gaps 2;  
QY 189 ACTAATCGATCTCTCGGTTGTGACAGATGACCTGGAAATTAATAAATACAAAACGAA 248  
DB 677 ACTAATCGGCTCTCTCGAATTTGTAAACAGATGACCTGGGAATTAATAAATTAAGACGAA 736  
QY 249 GTGCACGTGCTCATTTCAAGGACAGCCAAATAGATAATGAGACTTCGTTTCAATCATTT 308  
DB 737 GTGCACATGCTCATTTGAGGACAGCCAAATAGATAATGAGACTTCGTTTCAACATTT 796  
QY 309 TGCTACAGATGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCACC 368  
DB 797 TGCTACAGAAATGTAGCTGGGACCAATTTATATGTTTATGATGGGCACTCAATCTAOCGACC 856  
QY 369 GCTAGTGTGCTCAATTTAGTGGCTCATTTGCTCTGAGAGAGATGGCAATGAGACTGTCC 428  
DB 857 TCTGATTCGCTCTTTAGTGGCTCATTTGCTCTGAGAGAGATGGCAATGAGACGCTCC 916  
QY 429 TGAGTTGTTTGGCACATCAGTTATGCTTGTGCAATTTTATGATGGTGTCTTATAA 488  
DB 917 TGAGTCACTGTCACTTCAAGTTATGCACTGTGCAATTTTCAAGTATGCTGTATAA 976  
QY 489 TTTGACTGGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 548  
DB 977 TCTGACTGGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1036  
QY 549 AGGAGAGTGAAGTCAAGTAAATAGCAGCGATCTGTTGAATGGAATGTTTCTGAAACATG 608  
DB 1037 AGGAGAGTGAAGCAGATTAACAGCAGCAGCGCTGTTGAGTGAATGTTTCTGAAACATG 1096  
QY 609 GAAAGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCATCGAGG 668  
DB 1097 GAAAGGGGAGTGTGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCATCGAGG 1156  
QY 669 CATCTCAATTCAGTGTGTGAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCG 728  
DB 1157 CATCTGTAATGCAAGCGATACCAAGAGGTGCTCTGCTTCTCCTCACTGGCAGGGTCTCG 1216  
QY 729 ATGTTTCAGTTTCTGTATCCAGCTAACCACTCAATTTTGGACTCGAGAGGAATATTTCAACTT 788  
DB 1217 ATGTTCAATTTCTGTGCCAGCTAACCACTCAATTTTGGACTCGAGAGGAATATTTCTCATTT 1276  
QY 789 AAAGCTCCCGACAGCATCTCAATAAGCTGTGGTCAATGGAAACATTAATGTGGGTTGG 848  
DB 1277 AAAGCTTCCAGAGCTCTCATAAAGCTGTGGTCAATGGAAATTAATGTGGGTTGG 1336  
QY 849 AGGATATATGTTCAACCACTCAGATTAATAACATGTTTCTAGGATGATGACTTGTCTTAG 908  
DB 1337 CGATATATGTTTCAACCACTCAGATTAATAACATGTTTCTAGGATGATGACTTGTCTTAG 1396  
QY 909 GGAGTGGCTTCCACTAAACCGTTTCTGTGAACAATGTTGTTGTGTAGATATGGTCAATCTTT 968  
DB 1397 GGAATGGCTTCCACTAAACCGATTCGTGAACAGTGTGGTTGTGAAGATATGGTCAATCTTT 1456  
QY 969 GGCATTTATCAAGGATTAATAATTTACATGATGAGGAGAAATTTGATTCACCTGGGATGT 1028  
DB 1457 GGCATTTACATGAGTAAATTTACATGATGAGGAGAAATTTGATTCACAGGAGACGT 1516  
QY 1029 GACCAATGAGTTGAGAGTTTTCACATTCATATGAGTCAATGGGTTGTGTGACCCCTAA 1088  
DB 1517 GACCAATGAGCTGAGAGTATTTTCATATTCATATGAGTCAATGGGTTGTGTGACCCGAA 1576  
QY 1089 GGCAAAGGAGCAGTATGCAAGTGTGGCCTCTGCAACATGTTTACACTGAGAAATGG 1148

DB 1577 AGCTAAGGATCAGTATGAGTGGTGTGGACACTCAGCACACATTTGTTACCTGGCATCTGG 1636  
QY 1149 CCGAGTGTGTCATGCTGTGTCATCTTTGGTTCACCTGCTCTATGAGTATATAAGCAATGT 1208  
DB 1637 CCGTGTGTCATGTTTGGTTCATCTTCCGTCAATTTGCCACTCTATGAGTATATAAGCGTTGT 1696  
QY 1209 GCAGCAATATGATTTGGATAAGAACACATGGAGTATATTACACCCAGGGTGCCTTGT 1268  
DB 1697 GCAGCAATATGATTTGGATAAGAACACATGGAGTATATTACATCTCAGGGTGTCTTGT 1756  
QY 1269 GCAAGGGGTTACCGCCATAGCAGTGTTTACGACCATAGGACCAAGGCGCTTATAGTTCA 1328  
DB 1757 GCAAGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816  
QY 1329 TGGTGGCTTACAGGCTTTTCACTGTCCTTCACTGTCCTGAGTATGATCTCTACCGATA 1388  
DB 1817 TGGTGGCTTACAGGCTTTTCACTGTCCTTCACTGTCCTGAGTATGATCTCTACGATA 1876  
QY 1389 TGATGTGATACCCAGATGTGACCAATTTTAAAGCACAGCCGATTTTTCCTTACTTGCA 1448  
DB 1877 CGATGTGATACTCAGATGTGACCAATTTTAAAGCACAGCCGATTTTTCCTTACTTGCA 1936  
QY 1449 CACAGCTGTGATGAGTGAACCACTGCTGTGTTTGGGGGAAACACACAAATGACAC 1508  
DB 1937 TACAGCTGTGATGAGTGAACCACTGCTGTGTTTGGAGGGAAACACACAAATGACAC 1996  
QY 1509 ATCTATGAGCCATGGCGCCAAATGCTTCTTTCAGATTTTCATGGCCTATGACATTCCTG 1568  
DB 1997 TTCCATGAGCCACCGTGTGCAATGCTTCTCTCGGACTTTCATGGCTTATGACATTCCTG 2056  
QY 1569 TGACCGCTGTGCTGAGTGTTCCTCAGACCTGA-TTCCACCATGATGTCAACATTTGGCCA 1627  
DB 2057 TGACCGATGGTGTGCTTCCAGACCTGAGCTCCATCATGTGTCAACATTTGGCCA 2116  
QY 1628 TTCCAGCAGTCTTACACAAACAGCCATGATGTTGTCGGTGTTCATAGTCTCTCC 1687  
DB 2117 TT-CAGCAGTCTTGTACACAGCCTATGATGTTTCCGGCGCTTCAAGCCTCTCC 2175  
QY 1688 TCAGCGACATCCTGGTATTTACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGTT 1747  
DB 2176 TCAGTGCCTCTTGTCTTTTACCTCGGAGCAGTGGATGACACACCGCAGTGAAGCTGTT 2235  
QY 1748 GTTTAGCAGCAGGAGCTGTGATTTGGTGTGTGGAAACAGGGTGTCTCAGTGTATCT 1807  
DB 2236 GTGTGGCAGCAGGAGCTGTGATTCGGTGTCTGTGGGACACACAGTGTCTCGATGTACCT 2295  
QY 1808 CGTGGGCTCGCAACTGATGAACAAGAAAAAGTTAAAAATCAGAAATGTTTTTCCAAA 1867  
DB 2296 CTTGGAGTTGGCACTGAAGAAACAGCAGAAAGTTAAAAATCAGAGTGTTTTTCTAAA 2355  
QY 1868 GAACTTTGACCATGACAGATGTGACAGCAGCAGACAGATTTGTACAGCTGCACGCCAACA 1927  
DB 2356 GAACTTTGACCATGACAGATGTGACCCAGCAGCAGACAGATTTGTACAGCTGCACGCCAATA 2415  
QY 1928 CCAATGACTGCACTGGTGAATGACCATTTGTGTCCTCCAGGAAACACAGCTGTCTCAGAAG 1987  
DB 2416 CCAATGACTGCACTGGTGAATGATCAGTGTGTCCTGTGAAACACAGCTGCACAGAG 2475  
QY 1988 GCCAGATCTCCATTTTATGATGAGATTTGCCCAAGGATTAACCTATGATCTACTGTA 2047  
DB 2476 GCCAGATCTCCATTTGCCAAGTATGAGAGTTGCCCAAGGATTAACCTATGATCTACTGCA 2535  
QY 2048 ACAGAGAGCAGCTGAGGAGCTGTGCTGTGGACAGAACTGCGAGTGGGAGCCCGGA 2107  
DB 2536 ATAGAAACAGCTGAGGAGCTGTGCTGTAGCCAGAACTGCGAGTGGGAGCCCGGA 2595  
QY 2108 ATCAGGAGTCAATTTGCCCTGCCCAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGAA 2167  
DB 2596 ATCAAGAGTGCATCGCTTCCCGGAAATATCTGTGGCAATGGCTGCAATTTGGTTGAA 2655  
QY 2168 ACTCATGTTGAAAATTTACTCTCCAGAGGAGATTTATGACAAATGCTTAAATTTGTTGTA 2227  
DB 2656 ACTCGTGTGAAAATTTACTCTGCTAAGGAGATTTATGACAAATGCTTAAATTTGCTCTGTA 2715

2228 GGAACCAATGCGCTTTTGGCTTCTCTTACAAACCAGAAAGAGTAGAATTTGCTTTA 2287  
Db GGAACCAATGCGCTTTTGGCTTCTCTTACAAACCAGAAAGAGTAGAATTTGCTTTA 2775  
2288 AGCAGCTGCGAATAAGTCAGTCACTCTCAGAGCATGTCAGCTCACTTAACCCCATGG 2347  
Db AGCAGCTGCGAATAAGTCAGTCACTCTCAGAGCATGTCAGCTCACTTAACCCCATGG 2835  
2348 TCGGCTTCCGAAGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCATTTACAA 2407  
Db TCGGCTTCCGAAGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCATTTACAA 2895  
2408 ATAGTTTACAGTGTGATGCGCTGTGAGCCAGTGTGATGTCTGTGGAAATTTAT 2467  
Db ATAGTTTACAGTGTGATGCGCTGTGAGCCAGTGTGATGTCTGTGGAAATTTAT 2955  
2468 CAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTG 2527  
Db CAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTG 3015  
2528 TCTGTGAAGGCTTGCACCAACACAGTGTGAAGCAGTGCAGCAACCATGTGCTTGAGGA 2587  
Db TCTGTGAAGGCTTGCACCAACACAGTGTGAAGCAGTGCAGCAACCATGTGCTTGAGGA 3075  
2588 CAGCATGTGAGATTCACAGCGGCGAGCTCTGAGTGCATGTGTCAGCAACATCAAGC 2647  
Db CAGCATGTGAGATTCACAGCGGCGAGCTCTGAGTGCATGTGTCAGCAACATCAAGC 3135  
3076 CAGCGTGTGCGAGTGCACATAGCAGCAGCTCGGAGTGCATGTGTCAGCAACATCAAGC 3135  
2648 AGTGTGTGAGCTTCAATGCTATGTGCGCTTCCCTTTGCGCCAGTGTATGGAATGGT 2707  
Db AGTGTGTGAGCTTCAATGCTATGTGCGCTTCCCTTTGCGCCAGTGTATGGAATGGT 3195  
2708 ATAGCATGAGCACTGCGCCCTGGAATTTGTCAGCTACTGTACTGTAGTCAATGCT 2767  
Db ATAGCATGAGCACTGCGCCCTGGAATTTGTCAGCTACTGTACTGTAGTCAATGCT 3255  
3196 ATAGCATGAGCAGTGCACCTGGAATTTGCTGTGCTACTGTACTGTGAGCAATGCT 3255  
2768 TGGAGCAACCAAGGCTGTGGCTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGCA 2827  
Db TGGAGCAACCAAGGCTGTGGCTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGCA 3315  
3256 TGGAGCAGCCAGGCTGTGTGTGTGTACTGATCTAGCAATATCTGGCAAGGGAATGTA 3315  
2828 TAGAGGTTCTATAAAGGACCAAGTGAAGTGTCTGCAAGCCCTTCAAGAAATTTCT 2887  
Db TAGAGGTTCTATAAAGGACCAAGTGAAGTGTCTGCAAGCCCTTCAAGAAATTTCT 3375  
3316 TTGAGGCGAGCTATAAAGGACCTGTGAAGTGTGCGTCAAGGCTCTGCAAGAAATTTGT 3375  
2888 ATCCACAGCCCTCTCAATTTCCAGCATGTGTCTAGAGGACAGCATACAACTGTCTT 2947  
Db ATCCACAGCCCTCTCAATTTCCAGCATGTGTCTAGAGGACAGCATACAACTGTCTT 3435  
3376 ATCCACAGCCCTCTCTGAATCTCCAGCATGTGTCTAGAGGACAGCATACAACTGTCTT 3435  
2948 TCATTCTACTGTCCAGCTTGCACATGCAACGCGCACAGTAAATGCATCAATCAGAGCATCT 3007  
Db TCATTCTACTGTCCAGCTTGCACATGCAACGCGCACAGTAAATGCATCAATCAGAGCATCT 3495  
3436 TCATTCTACTGTCCAGCTTGCACATGCAACGCGCACAGTAAATGCATCAATCAGAGCATCT 3495  
3008 GTGAGAGTGTGAGAACTGACACAGCGCAAGCTGCGAGACTGATATCTGGCTTCT 3067  
Db GTGAGAGTGTGAGAACTGACACAGCGCAAGCTGCGAGACTGATATCTGGCTTCT 3555  
3496 GTGAGAGTGTGAGAACTGACACAGCGCAAGCTGCGAGACTGATATCTGGCTTCT 3555  
3068 ACCGTGTATCCCAATGAGGGAATGTCTAGCCATGCAAGTGCATGCGCAGCGCTCTC 3127  
Db ACCGTGTATCCCAATGAGGGAATGTCTAGCCATGCAAGTGCATGCGCAGCGCTCTC 3615  
3556 ATGTGTACCCGACTAATGAGGCAATGTGAGCCATGCAAGTGCATGCGCAGCGCTCTC 3615  
3128 TGTGAAACCAACACAGCGGCAAGTGTCTGACCAACAGGCGCTCAAGGGGAGGAGT 3187  
Db TGTGAAACCAACACAGCGGCAAGTGTCTGACCAACAGGCGCTCAAGGGGAGGAGT 3675  
3616 TGTGAAACCAACACAGCGGCAAGTGTCTGACCAACAGGCGCTCAAGGGGAGGAGT 3675  
3188 GCCAGCTATGTGAGTGAATAATGCATACCAAGGAACCTCTCAGAGGAACATGTTATT 3247  
Db GCCAGCTATGTGAGTGAATAATGCATACCAAGGAACCTCTCAGAGGAACATGTTATT 3735  
3676 GCCAGCTATGTGAGTGAATAATGCATACCAAGGAACCTCTCAGAGGAACATGTTATT 3735  
3248 ATACTTCTTATTGACTATCAGTTTACCTTTAGCTTATCCAGGAGATGATCGCTATT 3307  
Db ATACTTCTTATTGACTATCAGTTTACCTTTAGCTTATCCAGGAGATGATCGCTATT 3795  
3736 ATACCTTCTATTGACTATCAGTTTACCTTTAGCTTATCCAGGAGATGATCGCTATT 3795

3308 ACACAGCTATCAATTTTGTGCTACTCTGACGAACAAAACAGGATTTGGACATGTTCA 3367  
Db ACACAGCTATCAATTTTGTGCTACTCTGACGAACAAAACAGGATTTGGACATGTTCA 3855  
3368 TCAATGCTCCTCAAGAAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAA 3427  
Db TCAATGCTCCTCAAGAAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAA 3915  
3428 CCCAGGCTGGAGAGAGATGCTGTTTCAAAACCAACATTAAGGCTACAGAGATA 3487  
Db CCCAGGCTGGAGAGAGATGCTGTTTCAAAACCAACATTAAGGCTACAGAGATA 3975  
3488 GTTCTCTTAATGAGAAGTTTGATTTTCCCAACCAACATCACTTTCTTTTATG 3547  
Db GTTCTCTTAATGAGAAGTTTGATTTTCCCAACCAACATCACTTTCTTTTATG 4035  
3548 TCAGTAATTTCACTGGCCCATCAAAATTCAGGT 3581  
Db TCAGTAATTTCACTGGCCCATCAAAATTCAGGT 4069

RESULT 12  
AAS72660  
ID AAS72660 standard; cDNA; 3490 BP.  
XX  
AC AAS72660;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #8464.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR P-PSDB; ABG08473.  
XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 8464; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations

CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

Query Match 77.6%; Score 2790.4; DB 5; Length 3490;  
Best Local Similarity 94.2%; Pred. No. 0;  
Matches 3238; Conservative 0; Mismatches 51; Indels 147; Gaps 27;

QY 187 AGACTAACTGGATCTCTCTGGGTTGTGACAGATGGAACCTGGAATAATATAACAAACG 246  
DB 22 AGACTAACTGGATCTCTCTGGGTTGTGACAGATGGAACCTGGAATAATATAACAAACG 81

QY 247 AAGTGACAGTGGCTCATTTGAAGACAGACCAATAGAAATATGAGACTTCGTTTCAATCAT 306  
DB 82 AAGTGACAGTGGCTCATTTGAAGACAGACCAATAGAAATATGAGACTTCGTTTCAATCAT 141

QY 307 TTTGCTACAGAGTGTAGTTGGGACCAATTAATATGTTTATGATGGGACCTCAATTTATGCA 366  
DB 142 TTTGCTACAGAGTGTAGTTGGGACCAATTAATATGTTTATGATGGGACCTCAATTTATGCA 201

QY 367 CCCTAGTTGCTGCATTTAGTGGCTCATTTGCTGCTGATGCTGCTGCTGCTGCTGCTGCT 426  
DB 202 CCCTAGTTGCTGCATTTAGTGGCTCATTTGCTGCTGATGCTGCTGCTGCTGCTGCTGCT 261

QY 427 CCTGAGGTTGTTGCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 486  
DB 262 CCTGAGGTTGTTGCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321

QY 487 AATTTGACTGGATTTAATATTACTACAGTTTGTGATATGTCCTGCAATTAAGTCTGCTGCTG 546  
DB 322 AATTTGACTGGATTTAATATTACTACAGTTTGTGATATGTCCTGCAATTAAGTCTGCTGCTG 381

QY 547 CGAGGAGAGTGAAGATCAGTATAGCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606  
DB 382 CGAGGAGAGTGAAGATCAGTATAGCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 441

QY 607 TGAAGAGTGAAGATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 666  
DB 442 TGAAGAGTGAAGATGTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 501

QY 667 GGCATCTGCAATTCAGTGTGTCAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 726  
DB 502 GGCATCTGCAATTCAGTGTGTCAGAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561

QY 727 GATGTTTCAGTTCCTGTACCAGCTAACCCAGTCAATTTT-GGACTCGAGAGGAATATTTCTAA 785  
DB 562 GATGTTTCAGTTCCTGTACCAGCTAACCCAGTCAATTTTGGGACTCGAGAGGAATATTTCTAA 621

QY 786 CTT--AAAGTCCCGAGAGGATCTCAATAAGTGTGGTCAATGGAACATTAATGTGGTT 843  
DB 622 CTTTAAAGTCCCGAGAGGATCTCAATAAGTGTGGTCAATGGAACATTAATGTGGTT 681

QY 844 GTTGGAGGATATATGTTTCAACCACTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 903  
DB 682 GTTGGAGGATATATGTTTCAACCACTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCT 741

QY 904 TCTAGGAGTGGCTTCCACTAAACCGTCTGTGGAACAATGCTGGTGTGTTAGATGATGCTCAT 963  
DB 742 TCTAGGAGTGGCTTCCACTAAACCGTCTGTGGAACAATGCTGGTGTGTTAGATGATGCTCAT 801

QY 964 TCTTTGGCATTAACAAGGATTAATAATTACATGATGAGGAGGAAATTTGATTCACATGGG 1023  
DB 802 TCTTTGGCATTAACAAGGATTAATAATTACATGATGAGGAGGAAATTTGATTCACATGGG 861

QY 1024 AATGTGACCAATGAGTTGAGAGTTTTCATTAATGAGTCAATGAGTGTGTTGTTGCTGCTGCTG 1083  
DB 862 AATGTGACCAATGAGTTGAGAGTTTTCATTAATGAGTCAATGAGTGTGTTGTTGCTGCTGCTG 921

QY 1084 CCTAAGGCAAAAGGAGCAGTATGCACTGTGGGCACTCTGCACACATTTGTACACTGAAG 1143  
DB 922 CCTAAGGCAAAAGGAGCAGTATGCACTGTGGGCACTCTGCACACATTTGTACACTGAAG 981

QY 1144 AATGCCCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1203  
DB 982 AATGCCCGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041

QY 1204 AATGTGACAGGAATATGATTTTGATAAGAAACACATGAGGATATATTACACACCAGGGTGCC 1263  
DB 1042 AATGTGACAGGAATATGATTTTGATAAGAAACACATGAGGATATATTACACACCAGGGTGCC 1101

QY 1264 CTTGTGCAAGGGGTTTACGGCCATGAGCAGTGTTTACGACCATAGGACACAGGGCCCTATAC 1323  
DB 1102 CTTGTGCAAGGGGTTTACGGCCATGAGCAGTGTTTTACGACCATAGGACACAGGGCCCTATAC 1161

QY 1324 GTTCAATGCTGCTCTACAAGGCTTTTCAAGTCCCAATTAAGTACCGGCTTGCAGATGATCTCTAC 1383  
DB 1162 GTTCAATGCTGCTCTACAAGGCTTTTCAAGTCCCAATTAAGTACCGGCTTGCAGATGATCTCTAC 1221

QY 1384 CGATATGATGAGTATGCCAGATGTGACACCATTTTAAAGGACAGCCGATTTTTCGGTTAC 1443  
DB 1222 CGATATGATGAGTATGCCAGATGTGACACCATTTTAAAGGACAGCCGATTTTTCGGTTAC 1281

QY 1444 TTGACACAGCTGTGATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1503  
DB 1282 TTGACACAGCTGTGATAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1341

QY 1504 GACACATCTATGAGCCATGCGCCCAATGCTTCTTTCAGATTTTCATGGCCCTATGACATT 1563  
DB 1342 GACACATCTATGAGCCATGCGCCCAATGCTTCTTTCAGATTTTCATGGCCCTATGACATT 1401

QY 1564 GCCTGTGACCGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1622  
DB 1402 GCCTGTGACCGCTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1461

QY 1623 GGCATTTCCAGCAGCTTTACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1682  
DB 1462 GGCATTT-CAGCAGCTTTACACACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1520

QY 1683 CTTCTCTAGCAGCATTCCTGCTGATTTTCACTCGGAACAGTGTGATGCGGATCGGAGTGAAGC 1742  
DB 1521 CTTCTCTAGCAGCATTCCTGCTGATTTTCACTCGGAACAGTGTGATGCGGATCGGAGTGAAGC 1580

QY 1743 CGCTTGTGTTTACGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1802  
DB 1581 CGCTTGTGTTTACGACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1614

QY 1803 TATCTCGTGGCGCTGGCACTGATGAACAGAGAGAGGTTAAATATCAGAAATGTTTTC 1862  
DB 1615 ----- 1614

QY 1863 CAAAGAACTTTGACCATGACAGATGTGACACGACACAGATTTGTACAGCTGCACAGC 1922  
DB 1615 -----CTCTTGACCATGACAGATGTGACACGACACAGATTTGTACAGCTGCACAGC 1666

QY 1923 CAACCAATGACCTGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1982  
DB 1667 CAACCAATGACCTGCGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1726

QY 1983 AGAAGGCGAGATCTCCATTTTGTAGTATGAGATGTCGCCCAAGGATTAACCTTATGTACTA 2042  
DB 1727 AGAAGGCG-----AGGTATGAGATGTCGCCCAAGGATTAACCTTATGTACTA 1772

QY 2043 CTGTAAACAAGAGACACAGCTGCGGAGCTGTGCGCTGGAACAGAACTGCCAGTGGGAGCC 2102  
DB 1773 CTGTAAACAAGAGACACAGCTGCGGAGCTGTGCGCTGGAACAGAACTGCCAGTGGGAGCC 1832

QY 2103 CCGAATACAGAGTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGT 2162  
DB 1833 CCGAATACAGAGTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGT 1892

QY 2163 TGGAACTCATGTTTGAATAATTACTGCAAGGAGAAATTATGCAATGCTAAATGTT 2222  
 Db 1893 TGGAACTCATGTTTGAATAATTACTGCAAGGAGAAATTATGCAATGCTAAATGTT 1952  
 QY 2223 CTGTAGGAACCAATGCCCTTTTGGCTTCTCTTACAACCCAGAGAGGTA - GAATTTG 2281  
 Db 1953 CTGTAGGAACCAATGCCCTTTTGGCTTCTCTTACAACCCAGAGAGGTAAGGTTG 2012  
 QY 2282 TCCTTAAG--CAGTGGGAATATGCA-GTCATCTCAGA-GCATGTCCAA-GCTCACCTT 2336  
 Db 2013 TCCTTAAGGAGGCTGCGAATAATAGTGTGATCATCTCAGAGGCTGTCGAAGGCTCACCTT 2072  
 QY 2337 AACCCCATGG--TGGGCTCTCGGAAGATCAA-TGTGTCTACTGCTGCTGGAGATA 2392  
 Db 2073 AACCCCATGGGCTCGGCTCTCGGAAGATCAATTGTCTTACTGCTGGAGATA 2132  
 QY 2393 TGTCCCATTTACAATAGTTTACTACAGTGGATGCGTCTGAGCCAGTGATGCTGGAT 2452  
 Db 2133 TGTCCCATTTACAATAGTTTACTACAGTGGATGCGTCTGAGCCAGTGATGCTGGAT 2192  
 QY 2453 TCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2507  
 Db 2193 TCTGTGGAATTTTATTCAGAACCCAGTACTTCGGGGACTGAAGGCTGCAACCTGCAT 2252  
 QY 2508 --CAACCCACTCAATGTAGTGTCTGTGAAA--GGCCTGCAAAACCCAGT-GCTAAGCAG 2562  
 Db 2253 TCAACCCACTTCAATGTAGTGTCTGTGAAAAGGCTGTGCAAAACCCAGTGTGTAAGCAG 2312  
 QY 2563 TGGCGGACACATGTGCC-----TTGAGGACAGCATGTGGAGATGCAACGCGGAGCTC 2618  
 Db 2313 TGGCGGACACATGTGCCCTTTGAGGACAGCATGTGGAGATGCAACGCGGAGCTC 2372  
 QY 2619 TGACT-GCATGTGTGTCAGCAACATGAA-CCAGTGTGTGGACTCCCAATGCTATGTGGCC 2676  
 Db 2373 TGAGTGCATGTGTGTCAGCAACATGAAAGCAGTGTGTGGACTCCCAATGCTATGTGGCC 2432  
 QY 2677 TCCTTCCCTTTT-GGCCAGTGTATGGAATGGTATACATGA-CCACTGCCCCCTTGAAA 2734  
 Db 2433 TCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACATGAAGGACACTGCCCCCTTGAAA 2492  
 QY 2735 ATTGTTCAGGCTACTGTACCTGTAGTCTGATTCATTTGAGCAACCA-GCCTGTGGCTGGTGT 2793  
 Db 2493 ATTGTTCAGGCTACTGTACCTGTAGTCTGATTCATTTGAGCAACCAAGGCTGTGGCTGGTGT 2552  
 QY 2794 ACTGATCCAGCAATACTGCAAGGAAATGTCATAGAGGGTTCTCTATAAGGACCAAGTG 2853  
 Db 2553 ACTGATCCAGCAATACTGCAAGGAAATGTCATAGAGGGTTCTCTATAAGGACCAAGTG 2612  
 QY 2854 AAGATGCTTGGCAAGCCCTACAGGAATTTCTNTCCAGAGCCCTGCTCAATTCACG 2913  
 Db 2613 AAGATGCTTGGCAAGCCCTACAGGAATTTCTNTCCAGAGCCCTGCTCAATTCACG 2672  
 QY 2914 ATGTGTCTAGAGGACAGCAGATACAATCGG--TCTTTTCACTGTCTGAGCTTGCAT 2971  
 Db 2673 ATGTGTCTAGAGGACAGCAGATACAATCGGCTCTTTCATTTCACTGTCTGAGCTTGCAT 2732  
 QY 2972 GCAAC-GGCCACAGTAATGTCATCAATCAGAGCATCTGTGAGAGTGTGAGACCTGACC 3030  
 Db 2733 GGAACGGGCCACAGTAATGTCATCAATCAGAGCATCTGTGAGAGTGTGAGACCTGACC 2792  
 QY 3031 ACAGGCAAGCACTCGGAGACCTGCATATCTGGCTTCTAC--GGTGAATCCCAACCA-TGGAG 3088  
 Db 2793 ACAGGCAAGCACTCGGAGACCTGCATATCTGGCTTCTACGGGTGATCCCAACCAATGGAG 2852  
 QY 3089 GGAATGTACGCCATGCAAGTGCATGGCAGCGCTCTCTGTGCAACCAACCAACGCGCA 3148  
 Db 2853 GGAATGTACGCCATGCAAGTGCATGGCAGCGCTCTCTGTGCAACCAACCAACGCGCA 2912  
 QY 3149 AGTCTCTTCACCAACCAAGGCTGCAAGGGGACGAGTGCCAGCTATGTGAGGTAGAA 3208  
 Db 2913 AGTCTCTTCACCAACCAAGGCTGCAAGGGGACGAGTGCCAGCTATGTGAGGTAGAA 2972  
 QY 3209 ATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTTATTCATC 3268

Db 2973 ATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTTCTTATTGACTATC 3032  
 QY 3269 AGTTCACCTTTAGTCTCTATCCAGGAAGATGATGCTATTACACAGCTATCAATTTGTGG 3328  
 Db 3033 AGTTCACCTTTAGTCTCTATCCAGGAAGATGATGCTATTACACAGCTATCAATTTGTGG 3092  
 QY 3329 CTACTCTCTGACGAACAAACACAGGCAATTTGGACATGTTTCATCAATGCTCCAAAGAAATTCA 3388  
 Db 3093 CTACTCTCTGACGAACAAACACAGGATTTGGACATGTTTCATCAATGCTCCAAAGAAATTCA 3152  
 QY 3389 ACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTTGGAGAGAGATGC 3448  
 Db 3153 ACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTTGGAGAGAGATGC 3212  
 QY 3449 CTGTTGTTTCAAAACCAACCAATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG 3508  
 Db 3213 CTGTTGTTTCAAAACCAACCAATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTG 3272  
 QY 3509 ATTTTCGCAACCCCAATATCACTTTCTTTGTTTATGTCAGTAATTTCACTGGGCCCA 3568  
 Db 3273 ATTTTCGCAACCCCAATATCACTTTCTTTGTTTATGTCAGTAATTTCACTGGGCCCA 3332  
 QY 3569 TCAAAATTCAGGTGCA 3584  
 Db 3333 TCAAAATTCAGGTGAA 3348

RESULT 13  
 AA291922 standard; cDNA; 2625 BP.  
 ID AA291922 standard; cDNA; 2625 BP.  
 XX  
 AC AA291922;  
 XX  
 DT 08-JUN-2000 (first entry)  
 DE Human mahogany protein coding sequence #4.  
 XX  
 KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antibesity; anti-anorexic; anticachexic; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200005373-A2.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 21-JUL-1999; 99WO-US01.6484.  
 XX  
 PR 21-JUL-1998; 98US-003630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 99US-00245041.  
 XX  
 PA (MILL-) MILLENIUM PHARM INC.  
 XX  
 PI Moore K, Nagle DL;  
 XX  
 XX WPI; 2000-195103/17.  
 DR P-PSDB; AAY81809.  
 XX  
 PT New human and murine mahogany genes, useful, e.g. for diagnosis and  
 treatment of body weight disorders.  
 XX  
 PS Claim 1; Fig 20a; 188pp; English.  
 XX  
 CC This sequence represents a human mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (ii); (ii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-

CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (II) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (II) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (II)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX  
 SQ Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;

Query Match. 56.4%; Score 2028.8; DB 3; Length 2625;  
 Best Local Similarity 97.0%; Pred. No. 0;  
 Matches 2088; Conservative 0; Mismatches 62; Indels 2; Gaps 2;

QY 15 GCGCGCAACTGAGCGAAGCGCTGAGGAGGAGGAGCGCGCGGCGAGCGCTCGCGGCGAG 74  
 DB 234 GAGGCGGAGCGCGCTGCGGTGCGCGGCGCGGTGTCGCGCTCGCGCGAGCGGAGCGCAA 293  
 QY 75 GAGCGGCGGCGCGCACTGTGTCAACGCGGTGCGTGCACCCCTGCGACCGCGCGAGTGGCT 134  
 DB 294 GGAATGTGACCGCGCGTGTGTCAACGCGGTGCGTGCACCCCTGCGACCGCGCGAGTGGCT 353  
 QY 135 CTGCGCGCGCGCGTGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAAC 194  
 DB 354 CTGCGCGCGCGCGTGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAAC 413  
 QY 195 TGGATCTTCTGCGGTGTGTGACAGATGGACCTGGAAATTAATAACAAACGAAGTGCAC 254  
 DB 414 TGGATCTTCTGCGGTGTGTGACAGATGGACCTGGAAATTAATAACAAACGAAGTGCAC 473  
 QY 255 GTGGCTCATTTGAAGGACGCGCAATAGATATAGAGACTTGGTTTCAATCATTTTGTCTAC 314  
 DB 474 GTGGCTCATTTGAAGGACGCGCAATAGATATAGAGACTTGGTTTCAATCATTTTGTCTAC 533  
 QY 315 AGAGTGTAGTGGGACCATTTATATGTTATGATGGGACTCAATTTATGACCGGTAGT 374  
 DB 534 AGAGTGTAGTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGGTAGT 593  
 QY 375 TGCTGCATTTAGTGGGCTCATTTGCTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434  
 DB 594 TGCTGCATTTAGTGGGCTCATTTGCTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653  
 QY 435 TGTTCGCACATCAGGTATGCTTGTGCTGCAATTTTATGATGCTGCTTAAATTTGAC 494  
 DB 654 TGTTCGCACATCAGGTATGCTTGTGCTGCAATTTTATGATGCTGCTTAAATTTGAC 713  
 QY 495 TGGATTTAATTTACTTACAGTTTGTGATGTGCTCCAAATAACTGCTCAGGCGGAGGAGA 554  
 DB 714 TGGATTTAATTTACTTACAGTTTGTGATGTGCTCCAAATAACTGCTCAGGCGGAGGAGA 773  
 QY 555 GTGTAGATCAGTAAATAGAGCGATGCTTGTGAATGTGAATGTTCTTGAAATCTGGAAGG 614  
 DB 774 GTGTAGATCAGTAAATAGAGCGATGCTTGTGAATGTGAATGTTCTTGAAATCTGGAAGG 833  
 QY 615 TGAAGCATGTGACATTTCTCAGTGTACAGCAACTGTGGTTTTCCTCATCGGCGATCTG 674  
 DB 834 TGAAGCATGTGACATTTCTCAGTGTACAGCAACTGTGGTTTTCCTCATCGGCGATCTG 893  
 QY 675 CAATTCAGTGTGTGAGAGGATGCTTCCTGCTCTCTCAGACTGGCAGGGTCTGGATGTTTC 734  
 DB 894 CAATTCAGTGTGTGAGAGGATGCTTCCTGCTCTCTCAGACTGGCAGGGTCTGGATGTTTC 953  
 QY 735 AGTTCCTGTACAGCTAACAGGTCATTTTGGACTCGAGGAGATATTTCTAACTTAAGCT 794  
 DB 954 AGTTCCTGTACAGCTAACAGGTCATTTTGGACTCGAGGAGATATTTCTAACTTAAGCT 1013  
 QY 795 CCOCAGGACATCTCAATAAGCTGTGGTCAATGAAACATTTATGCGGTTGTTGGAGGATA 854

DB 1014 CCCAGAGCATCTCATAAAGCTGTGGTCAATGGAACAAANTATGTGGGTGTTGGAGGATA 1073  
 QY 855 TATGTTCAACCACTCAGATTAATAACATGGTTCTTAGCGTATGACCTTGTCTTAGGAGTG 914  
 DB 1074 TATGTTCAACCACTCAGATTAATAACATGGTTCTTAGCGTATGACCTTGTCTTAGGAGTG 1133  
 QY 915 GCTTCCACTAAACCGTCTCTGGAACATGTGGTTGTTAGATATGCTATTCCTTGGCAAT 974  
 DB 1134 GCTTCCACTAAACCGTCTCTGGAACATGTGGTTGTTAGATATGCTATTCCTTGGCAAT 1193  
 QY 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAAATGATTCAACTGGGAATGTGACCAA 1034  
 DB 1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAAATGATTCAACTGGGAATGTGACCAA 1253  
 QY 1035 TGAGTTGAGAGTTTTTTCATTTCAATTAATAGACTCATGGGTGTTGTGACCCCTTAAGGCAA 1094  
 DB 1254 TGAGTTGAGAGTTTTTTCATTTCAATTAATAGACTCATGGGTGTTGTGACCCCTTAAGGCAA 1313  
 QY 1095 GGAGCAGTATGCAAGTGGTGGGCACTCTGCAACATTTGTACACTGAAGAAATGCCCGAGT 1154  
 DB 1314 GGAGCAGTATGCAAGTGGTGGGCACTCTGCAACATTTGTACACTGAAGAAATGCCCGAGT 1373  
 QY 1155 GGTCAATGCTGTCTATCTTTTGGTCACTGCCCTCTCTATGGAATATAAGCAATGTGACGGA 1214  
 DB 1374 GGTCAATGCTGTCTATCTTTTGGTCACTGCCCTCTCTATGGAATATAAGCAATGTGACGGA 1433  
 QY 1215 ATATGATTTGATTAAGACACATGAGTATATTACACACCCAGGTCGCCCTTGTGCAAGG 1274  
 DB 1434 ATATGATTTGATTAAGACACATGAGTATATTACACACCCAGGTCGCCCTTGTGCAAGG 1493  
 QY 1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACAGGCGCCCTATACGTTCAATGTTGG 1334  
 DB 1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACAGGCGCCCTATACGTTCAATGTTGG 1553  
 QY 1335 CTACAGGCTTTTCAAGTGCATTAAGTACCGGCTTTCAGATGATCTCTACCCATATGATGT 1394  
 DB 1554 CTACAGGCTTTTCAAGTGCATTAAGTACCGGCTTTCAGATGATCTCTACCCATATGATGT 1613  
 QY 1395 GGATACCCAGATGTGGACCATTTTAAGGACAGCGCGATTTTCCGTTACTTGGCACACAGC 1454  
 DB 1614 GGATACCCAGATGTGGACCATTTTAAGGACAGCGCGATTTTCCGTTACTTGGCACACAGC 1673  
 QY 1455 TGTGATGATGAGTGAACCATGCTGGTGTGTTGGGGGAAAACACACAATGACATCTAT 1514  
 DB 1674 TGTGATGATGAGTGAACCATGCTGGTGTGTTGGGGGAAAACACACAATGACATCTAT 1733  
 QY 1515 GAGCCATGCGCCCAATGCTTCTCTCAGATTTTCAGCTTACGCTATGACATGCTGACCG 1574  
 DB 1734 GAGCCATGCGCCCAATGCTTCTCTCAGATTTTCAGATTTTCATGCTTATGACATGCTGACCG 1793  
 QY 1575 CTGGTCACTGCTTCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTTCCAG 1633  
 DB 1794 CTGGTCACTGCTTCCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTT-CAG 1852  
 QY 1634 CAGTCTTACACACAGACCACTGATGTGTTGCGTGGTTCATAGTCTCTCTCTCAGCG 1693  
 DB 1853 CAGTCTTACACACAGACCACTGATGTGTTGCGTGGTTCATAGTCTCTCTCTCAGCG 1912  
 QY 1694 ACATCTCTGATTTCACTTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCCTTGTGTAG 1753  
 DB 1913 ACATCTCTGATTTCACTTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCCTTGTGTAG 1792  
 QY 1754 CAGCAGGACCTTGGTATTTGCGGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTGTTGG 1813  
 DB 1973 CAGCAGGACCTTGGTATTTGCGGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCTGTTGG 2032  
 QY 1814 CGCTGGCAACTGTGAACAGAAAGTTPAAATCAGATGTTTTCAAAAAGAACTC 1873  
 DB 2033 CGCTGGCAACTGTGAACAGAAAGTTPAAATCAGATGTTTTCAAAAAGAACTC 2092  
 QY 1874 TTGACCATGACAGATGTGACCCAGACACAGATTTGTTTACAGCTGACACCAACCAATG 1933  
 DB 2093 TTGACCATGACAGATGTGACCCAGACACAGATTTGTTTACAGCTGACACCAACCAATG 2152



QY 1934 ACTGCCACTGGTCAATGACCATTTGTCCTCCAGGAAACACACCTGCTCAGAGGCCAGA 1993  
 Db 2153 ACTGCCACTGGTCAATGACCATTTGTCCTCCAGGAAACACACCTGCTCAGAGGCCAGA 2212  
 QY 1994 TCTCCATTTTATGATATGAGAAATGCCCCAAGGATAACCTATGTAATCTACTACTGTAAACAAG 2053  
 Db 2213 TCTCCATTTTATGATATGAGAAATGCCCCAAGGATAACCTATGTAATCTACTACTGTAAACAAG 2272  
 QY 2054 AGACCACTGACAGAGCTGTCCTGGAACAGAACTGCCAGTGGAGCCCCCGGAATCAGG 2113  
 Db 2273 AGACCACTGACAGAGCTGTCCTGGAACAGAACTGCCAGTGGAGCCCCCGGAATCAGG 2332  
 QY 2114 AGTGCAATGCCCTGCCGAAATATCTGTGGCAATGGCTGGCATTTGGTTGG 2165  
 Db 2333 AGTGCAATGCCCTGCCGAGGCTGAGGCTGCAAGGCTCATCTTGGTGTGTGG 2384

## RESULT 14

AAZ91917

ID AAZ91917 standard; cDNA; 2419 BP.

XX AAZ91917;

AC AAZ91917;

XX AAZ91917;

DT 08-JUN-2000 (first entry)

XX Murine mahogany protein coding sequence akml003.  
 XX Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
 XX Mus sp.  
 OS Mus sp.  
 XX WO200005373-A2.  
 PN 03-FEB-2000.  
 XX 21-JUL-1999; 99WO-US016484.  
 XX 21-JUL-1998; 98US-0093630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 98US-00245041.  
 XX (MILL-) MILLENIUM PHARM INC.  
 PA Moore K, Nagle DL;  
 PI WPI; 2000-195103/17.  
 DR P-PSDB; AAY81804.  
 XX

PT New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.  
 XX Claim 1; Fig 8a; 188pp; English.

XX This sequence represents a murine mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iv) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (ii) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (ii)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other

CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate

XX Sequence 2419 BP; 578 A; 567 C; 680 G; 594 T; 0 U; 0 Other;  
 SQ

Query Match 39.7%; Score 1428.8; DB 3; Length 2419;  
 Best Local Similarity 89.4%; Pred. No. 0;  
 Matches 1560; Conservative 0; Mismatches 182; Indels 2; Gaps 2;

QY 199 ACTAACTGGATCTTCTGGTTTGCACAGATGGACCTGGAAATATATAAATACAAAGCAA 248

Db 677 ACTAACTGGCTCTTCTGGATTGTAACAGATGGACCTGGGAAATATATAAATACAAAGCAA 736

QY 249 GTGCACGTGGCTCATTTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTT 308

Db 737 GTGCACATGGCTCATTTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTT 796

QY 309 TGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCACC 368

Db 797 TGCTACAGATGTAGTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCACC 856

QY 369 GCTAGTTGCTGCATTTAGTGGCTCATTTGTTCTTGCAGAGATGGCAATCAGACTGTCCC 428

Db 857 TCTGATTCCTGCTTTAGTGGCTCATTTGTTCTTGCAGAGATGGCAATCAGACTGTCCC 916

QY 429 TGAGTTGTTGCCACATCAGGTTATGCTTGTGCTCATTTTATGATGCTGCTTATAA 488

Db 917 TGAGTCACTGCTCATTTAGGTTATGCTGCTCATTTTATGATGCTGCTTATAA 976

QY 489 TTTGACTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 548

Db 977 TCTGACTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1036

QY 549 AGGAGAGTGAAGATCAAGTAAATAGCAGCATCTGTTGTAATGTGAATGTCTGAAACTG 608

Db 1037 AGGAGAGTGAAGAGCAGTAAACAGCAGCGCTGTTGAGTGTGAATGTCTGAAACTG 1096

QY 609 GAAAGGTGAGCATGTGACATTCCTCACTGTACAGACAACCTGTTGTTTCTCATCAGG 668

Db 1097 GAAAGGGGAGTGGTGTGACATTCCTCACTGTACAGACAACCTGTTGTTTCTCATCAGG 1156

QY 669 CATCTGCAATCAAGTGTATGTCAGAGATGCTCTCTGCTTCTCAGACTGGCAGGCTCTGG 728

Db 1157 CATCTGTAATGCAAGCATACAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1216

QY 729 ATGTTCACTTCTGTACAGTAACAGTCAATTTTGGACTCGAGAGAAATATCTTAATT 788

Db 1217 ATGTTCAATTCCTGTGCGAGCTAAACCACTGTTTGGACTCGAGAGAAATATCTTAATT 1276

QY 789 AAAGCTCCAGAGCATCTCATAAAGCTGTGCTCAATGGAACATTAATGTTGGTGTGTTGG 848

Db 1277 AAAGCTCCAGAGCATCTCATAAAGCTGTGCTCAATGGAACATTAATGTTGGTGTGTTGG 1336

QY 849 AGGATATATGTTCAACCACTCAGATTATAACAGTCTCTAGCGTATGACCTTCTCTAG 908

Db 1337 CGGATATATGTTCAACCACTCAGATTATAACAGTCTCTAGCGTATGACCTTCTCTAG 1396

QY 909 GGAGTGGCTTCCACTAAACCTGCTGTCAACATGCTGTTGTTAGATATGCTCATCTTT 968

Db 1397 GGAGTGGCTTCCACTAAACCTGCTGTCAACATGCTGTTGTTAGATATGCTCATCTTT 1456

QY 969 GGCATTATACAGGATATAATTTATCATGTATGAGGAAAAATGATTCAACTGGGAATGT 1028

Db 1457 GGCATTATACAGGATATAATTTATCATGTATGAGGAAAAATGATTCAACTGGGAATGT 1516

QY 1029 GACCAATGAGTGTAGATTTTTCATCTCAATTAAGTGTGCTGTTGTTGTTGTTGTTGTTG 1088

Db 1517 GACCAATGAGTGTAGATTTTTCATCTCAATTAAGTGTGCTGTTGTTGTTGTTGTTGTTG 1576

QY 1089 GGCAGAGGAGCATGTGAGTGGTGGGCACTCTGCACACATTTGTACACTGAAGAAATGG 1148

Db 1577 AGCTAAGGATCAGTATGAGTGGTGGGCACTCTGCACACATTTGTACACTGAAGAAATGG 1636



QY 1149 CCGAGTGGTCACTGCTGGTCACTCTTTGGTCACTGCTCCCTCTCTATGGATATATAAGCAATGT 1208  
 DB |||||  
 QY 1637 CCGTGTGGTCACTGCTGGTCACTCTTTGGTCACTGCTCCCTCTCTATGGATATATAAGCAATGT 1696  
 DB |||||  
 QY 1209 GCAGGAATATGATTTGGATTAAGAACACATGGAGTATATATACACACCCAGGGTGCCTTGT 1268  
 DB |||||  
 QY 1697 GCAGGAATATGATTTGGATTAAGAACACATGGAGTATATATACACACCCAGGGTGCCTTGT 1756  
 DB |||||  
 QY 1269 GCAAGGGGGTTACGGGCTAGCAGTGTATACGACCATAGGACACAGGGCCCTATAGCTTCA 1328  
 DB |||||  
 QY 1757 GCAAGGGGGTTATGGGCAAGTAGTGTATATGATGACAGCAAGGCTCTAGCTTCA 1816  
 DB |||||  
 QY 1329 TGGTGGCTCAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTTACCGATA 1388  
 DB |||||  
 QY 1817 TGGTGGCTCAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTTACCGATA 1876  
 DB |||||  
 QY 1389 TGATGTGGATACCCAGATGACCATCTTAAAGGACAGCCGATTTTCGGTACTTGA 1448  
 DB |||||  
 QY 1877 CGATGTGGATACCTCAGATGTGGACCAATCTTAAAGGACAGCCGATTTTCGGTACTTGA 1936  
 DB |||||  
 QY 1449 CACAGCTGTGATAGTGGAGAACCATGCTGGTGTGGGGGAAACACACACAATGACAC 1508  
 DB |||||  
 QY 1937 TACAGCTGTGATAGTGGAGAACCATGCTGGTGTGGGGGAAACACACACAATGACAC 1996  
 DB |||||  
 QY 1509 ATCTATGAGCCATGGCGCCCAATGCTTCTTCAGATTTTCAGGCTATGACATTCGCTG 1568  
 DB |||||  
 QY 1997 TTCATGAGCCATGGCGCCCAATGCTTCTTCAGATTTTCAGGCTATGACATTCGCTG 2056  
 DB |||||  
 QY 1569 TGACCGCTGTGATGCTTCCAGACCTGA-TTCCACCATGATGTCAACAGATTTGGCCA 1627  
 DB |||||  
 QY 2057 TGACCGATGTGATGCTTCCAGACCTGAGCTCCATCATGATGCAACAGATTTGGCCA 2116  
 DB |||||  
 QY 1628 TTCACGAGTCTTACACACAGCACCATGATGATGTTGGTGGTTCATATGCTCTCTCC 1687  
 DB |||||  
 QY 2117 TT-CAGCAGTCTTGTACACAGCACCATGATGATGTTGGTGGTTCATATGCTCTCTCC 2175  
 DB |||||  
 QY 1688 TCAGGCACATCTGCTGATTCACCTCGGACAGTGTGATGCTGATGCTGATGAGCGCTT 1747  
 DB |||||  
 QY 2176 TCAGTACGCTTGTGCTTTTACCTCGAGCAGTGTGATGCTGATGCTGATGAGCGCTT 2235  
 DB |||||  
 QY 1748 GTTTAGCAGCAGCAGCTGCTGATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1807  
 DB |||||  
 QY 2236 GTTGGCAGCAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2295  
 DB |||||  
 QY 1808 CTTGGGCGCTGGCACTGATGACACAGAGAGAGAGTAAATCAGATCTTTTTCMAAA 1867  
 DB |||||  
 QY 2296 CTTGGGAGTTGGCACTGAGAACACAGAGAGAGAGTAAATCAGATCTTTTTCMAAA 2355  
 DB |||||  
 QY 1868 GAACTCTTGACCATGACAGATGTGACAGCACACAGATTTGTACAGCTGCACAGCCACA 1927  
 DB |||||  
 QY 2356 GAACCTTTGACCATGACAGATGTGACAGCACACAGATTTGTACAGCTGCACAGCCACA 2415  
 DB |||||  
 QY 1928 CCAA 1931  
 DB |||||  
 QY 2416 CCAA 2419

## RESULT 15

AAZ91919 standard; cdNA; 6373 BP.

QY 2035 ATGTACTACTGTAAACAGAACAGCAGCTGACGAGAGCTGTGCCCTGGACACCAATGCCAG 2094  
 DB |||||  
 QY 2095 TGGGAGCCCCGGATCAGAGTGCATTGCCCTGCCGAAAATATCTGTGGCATTTGGCTGG 2154  
 DB |||||  
 QY 2155 CATTGTGTGAAACTCATGTTTGAATAATTACTACTGCCAAGGAGAAATTATGCAATGCT 2214  
 DB |||||  
 QY 2215 AAATGTCTCTAGACCAACCAATGCCCTTTTGGCTTCTTCTTACACCCCAAGAGGTA 2274  
 DB |||||

Query Match 37.3%; Score 1341.6; DB 3; Length 6373;  
 Best Local Similarity 95.2%; Pred. No. 0;  
 Matches 1498; Conservative 7; Mismatches 42; Indels 26; Gaps 11;

QY 2035 ATGTACTACTGTAAACAGAACAGCAGCTGACGAGAGCTGTGCCCTGGACACCAATGCCAG 2094  
 DB |||||  
 QY 2095 TGGGAGCCCCGGATCAGAGTGCATTGCCCTGCCGAAAATATCTGTGGCATTTGGCTGG 2154  
 DB |||||  
 QY 2155 CATTGTGTGAAACTCATGTTTGAATAATTACTACTGCCAAGGAGAAATTATGCAATGCT 2214  
 DB |||||  
 QY 2215 AAATGTCTCTAGACCAACCAATGCCCTTTTGGCTTCTTCTTACACCCCAAGAGGTA 2274  
 DB |||||

Human mahogany protein coding sequence #1.  
 Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
 weight regulation; cell therapy; body weight disorder; cachexia;  
 anorexia; hyperpigmentation; increased metabolic rate disorder;  
 hyperphagia; Antibesia; antianorexic; anticachexic; ss.  
 Homo sapiens.

Key Location/Qualifiers  
 misc\_difference 448  
 /tag= a  
 /note= "represented in specification as :"  
 misc\_difference 615  
 /tag= b  
 /note= "represented in specification as :"  
 misc\_difference 683  
 /tag= c  
 /note= "represented in specification as :"

WO200005373-A2.

03-FEB-2000.

21-JUL-1999; 99WO-US016484.

21-JUL-1998; 98US-0093630P.

20-OCT-1998; 98US-0104978P.

05-FEB-1999; 99US-00245041.

(MILL-) MILLENIUM PHARM INC.

Moore K, Nagle DL;

WPI; 2000-195103/17.

P-PSDB; AAY81806.

New human and murine mahogany genes, useful, e.g. for diagnosis and treatment of body weight disorders.

Claim 1; Fig 10a; 188pp; English.

This sequence represents a human mahogany gene of the invention. The mahogany genes are used: (i) to produce recombinant mahogany (mg) proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming therapeutics; (iii) as a source of diagnostic probes and primers for detecting expression of mg genes or mutations, regulatory defects, in this gene, or for isolation of related sequences; and (iv) in (cell-based) gene therapy. (ii) are used to raise specific antibodies (Ab); to identify other (extra)cellular products involved in weight regulation, and to screen for agents that disrupt interaction between (ii) and other macromolecules. The Ab are used to detect abnormal levels (or function) of (ii) (for diagnosis, prognosis or monitoring of treatment); to evaluate (ii)-expressing cells intended for cell therapy, and as therapeutic mg inhibitors. Cells that express the mg gene (or contain the mg polypeptide) are used to identify agents (A) that modulate mg activity. (A) are potentially useful for the treatment of body weight disorders, particularly obesity, cachexia or anorexia, or other conditions associated with the mg gene such as hyperpigmentation, hyperphagia and disorders that result in increased metabolic rate

Sequence 6373 BP; 1736 A; 1468 C; 1460 G; 1699 T; 0 U; 10 Other;

Db 181 AAATGTTCTGTAGAACCAACAATGCCCTTTGGCTTCTCTTACAAACCCAGAGAGGTA 240  
Qy 2275 GAATTTGCTCTTAAGCAGCTGCGAATAATGCAATGATCTCAGAGCATGCTCCAGCTCAAC 2334  
Db 241 GAATTTGCTCTTAAGCAGCTGCGAATAATGCAATGATCTCAGAGCATGCTCCAGCTCAAC 300  
Qy 2335 TTAACCCCATGGGTGCGCCCTT---CGGAAGATCAATGTGTCTTACTG---TGCTGGG 2386  
Db 301 TTAACCCCATGGGTGCGCCCTTGGGAAGGTYCAATGTGTCTTACTKGTGTCTGGGAAG 360  
Qy 2387 AAGATATGTCCTCCATTTACAAATAG---TTTACTACAGTGGATGCCGTCTGA---GCCAG 2441  
Db 361 GATATGTCCTCCATTTTACAAATAGTTTACTACAGTGGATGCCGTCTTGAAGCCCAAG 420  
Qy 2442 TGATGCT--GGAATCTGTGAAATTTTATC---AGAACCCAGT--ACTCGGGACTGAAG 2493  
Db 421 TGTGCTTGGATTCGTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGACTGAAG 480  
Qy 2494 GCTGCAACCTGCATCAACCC--ACTCAATGGTAGTGTCTGTGAAGCCCTGCNAACCA 2551  
Db 481 GCTGCAACCTGCATCAACCCCACTYMAATGGTAGTGTCTGTGAAGCCCTGCNAACCA 540  
Qy 2552 GTGCTAA--GCAGTGCCTCCGACACCATGTGCTTGAAGCAGCATGTGGAGATTGCACAGC 2610  
Db 541 GTGCTAAGCAGTGCCTCCGACACCATGTGCTTGAAGCAGCATGTGGAGATTGCACAGC 600  
Qy 2611 GGCAGCTCTGAGTG--CATGTGTGCGACAAATGAAGCAGTGTGTGGATCTCCATGCCCTA 2669  
Db 601 GGCAGCTCTGAGTGCATGTGTGTGCGACAAATGAAGCAGTGTGTGGATCTCCATGCCCTA 660  
Qy 2670 TGTGGCTCTCTCCCTTTGCG--CCAGTGTATGAATGGTATACGATGAGCAGCTGCCCC 2728  
Db 661 TGTGGCTCTCTCCCTTTGCGNCAGTGTATGAATGGTATACGATGAGCAGCTGCCCC 720  
Qy 2729 CTGAAATTTGTCAGGCTACTGACCTGTAGTCAATTCCTTGGAGCAACAGGCTGTGGCT 2788  
Db 721 CTGAAATTTGTCAGGCTACTGACCTGTAGTCAATTCCTTGGAGCAACAGGCTGTGGCT 780  
Qy 2789 GGTGTACTGTATCCAGCAATACTGGCAAGGGAATGATAGAGGCTTCTTATTAAGGAC 2848  
Db 781 GGTGTACTGTATCCAGCAATACTGGCAAGGGAATGATAGAGGCTTCTTATTAAGGAC 840  
Qy 2849 CAGTGAAGATGCCTTTCAGAGCCCTTACAGGAATTTCTATCCAGCCCTGCTCAATT 2908  
Db 841 CAGTGAAGATGCCTTTCAGAGCCCTTACAGGAATTTCTATCCAGCCCTGCTCAATT 900  
Qy 2909 CCAGCATGTCTTAGAGGACAGCAGATACAACTGTGCTTTTCAITCACTGTCCAGCTGCC 2968  
Db 901 CCAGCATGTCTTAGAGGACAGCAGATACAACTGTGCTTTTCAITCACTGTCCAGCTGCC 960  
Qy 2969 AATGCAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGA 3028  
Db 961 AATGCAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTGA 1020  
Qy 3029 CCACAGCAAGCACTGCGACCTGCAATCTGGCTTTCTACGGTGTATCCCAATGGAG 3088  
Db 1021 CCACAGCAAGCACTGCGAGACCTGCAATCTGGCTTTCTACGGTGTATCCCAATGGAG 1080  
Qy 3089 GGAATGTGACCATGCAAGTGAATGGGACGCGCTCTGTGCAACCAACACCGGCA 3148  
Db 1081 GGAATGTGACCATGCAAGTGAATGGGACGCGCTCTGTGCAACCAACACCGGCA 1140  
Qy 3149 AGTGTCTTTCACCAACCAAGGGGTCAAGGGGACGAGTGCAGCTATGTAGGTTAGAAA 3208  
Db 1141 AGTGTCTTTCACCAACCAAGGGGTCAAGGGGACGAGTGCAGCTATGTAGGTTAGAAA 1200  
Qy 3209 ATCGATACCAAGGAACCTCTCAGAGGACATGTTATATCTCTTCTTATGACTATC 3268  
Db 1201 ATCGATACCAAGGAACCTCTCAGAGGACATGTTATATCTCTTCTTATGACTATC 1260  
Qy 3269 AGTTACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAATTTTGTGG 3328  
Db 1261 AGTTACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCAATTTTGTGG 1320

Qy 3329 CTACTCTTGACGAACAAACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTCA 3388  
Db 1321 CTACTCTTGACGAACAAACAGGATTTGGACATGTTTCATCAATGCTTCCAGAAATTCA 1380  
Qy 3389 ACCTCAACATCACCCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGC 3448  
Db 1381 ACCTCAACATCACCCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGATGC 1440  
Qy 3449 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATCAGAAGTTTG 3508  
Db 1441 CTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATCAGAAGTTTG 1500  
Qy 3509 ATTTTGGCAACCAACCAAAATATCACTTTCTTTGTTTATGTCAGTAATTTACCTGGCCCA 3568  
Db 1501 ATTTTGGCAACCAACCAAAATATCACTTTCTTTGTTTATGTCAGTAATTTACCTGGCCCA 1560  
Qy 3569 TCAAAATTCAGGT 3581  
Db 1561 TCAAAATTCAGAT 1573

Search completed: March 4, 2004, 18:08:57  
Job time : 860.6 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	3494.4	97.1	4072	9	US-09-893-238-16		Sequence 16, Appl
2	3476.8	96.1	8589	9	US-09-893-238-14		Sequence 14, Appl
3	2845.2	79.1	8827	9	US-09-893-238-11		Sequence 11, Appl
4	2028.8	56.4	2625	9	US-09-893-238-18		Sequence 18, Appl
5	1428.8	39.7	2419	9	US-09-893-238-8		Sequence 8, Appl
6	1364.4	37.9	6370	9	US-09-893-238-12		Sequence 12, Appl
7	1200.3	33.4	6733	15	US-10-197-824-1		Sequence 1, Appl
8	463	12.9	1051	9	US-09-893-238-10		Sequence 10, Appl
C	277.4	7.7	625	14	US-10-358-584-5869		Sequence 5869, Ap
	250.8	7.0	531	9	US-09-884-761-15900		Sequence 15900, A
10	250.8	6.8	246	9	US-09-736-457-966		Sequence 966, App
C 11	246	6.8	246	9	US-09-902-941-966		Sequence 966, App
C 12	246	6.8	246	9	US-09-849-626-966		Sequence 966, App
C 13	246	6.8	246	9	US-10-017-754-966		Sequence 966, App
C 14	246	6.8	246	14	US-10-113-872-966		Sequence 966, App
C 15	246	6.8	246	14	US-10-113-872-966		Sequence 966, App

Qy	135	CTGCCCGCCGGCTGGGTGGCGCAGCAAAATGCCAGCATCTGGGGGGCCGCTTTCAGACTAAC	194
Db	354	CTGCCCGCCGGCTGGGTGGCGCAGCAAAATGCCAGCATCTGGGGGGCCGCTTTCAGACTAAC	413
Qy	195	TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTTATAAATACAAAACGAAGTCAC	254
Db	414	TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTTATAAATACAAAACGAAGTCAC	473
Qy	255	GTGGCTCAATTGAAGGACAGCCAAAATAGAATAATCAGACTTCGTGTTTCAATCATTTTGTCTAC	314
Db	474	GTGGCTCAATTGAAGGACAGCCAAAATAGAATAATCAGACTTCGTGTTTCAATCATTTTGTCTAC	533
Qy	315	AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACATCAATTTATGCACCGCTAGT	374
Db	534	AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACATCAATTTATGCACCGCTAGT	593
Qy	375	TGCTGCAATTTAGTGGCCCTCAATGTTTCCTGAGAGAGATGGCAATCAGACTGTCCCTGAGGT	434
Db	594	TGCTGCAATTTAGTGGCCCTCAATGTTTCCTGAGAGAGATGGCAATCAGACTGTCCCTGAGGT	653
Qy	435	TGTTGCCACATCAGGTTATGCTTGTCTGCAATTTTTTTTAGTGAATGCTGCTTTATAATTTCAC	494
Db	654	TGTTGCCACATCAGGTTATGCTTGTCTGCAATTTTTTTTAGTGAATGCTGCTTTATAATTTCAC	713
Qy	495	TGGATTTAAATTAATCTTTACAGTTTTGATGTGTGCCAAATAACTGCTCAGGCCGAGCAGA	554
Db	714	TGGATTTAAATTAATCTTTACAGTTTTGATGTGTGCCAAATAACTGCTCAGGCCGAGCAGA	773
Qy	555	GTGTAAGATCAGTAATAAGACGCGATCTGTTGAAATGTAATGTTTCTGAAAACCTGGAAGG	614
Db	774	GTGTAAGATCAGTAATAAGACGCGATCTGTTGAAATGTAATGTTTCTGAAAACCTGGAAGG	833
Qy	615	TGAAGCATGTCAACATTCCTCACTGTACAGACAACTGTGTTTTCTCATCGAGGCATCTG	674
Db	834	TGAAGCATGTCAACATTCCTCACTGTACAGACAACTGTGTTTTCTCATCGAGGCATCTG	893
Qy	675	CAATTCAAGTGATCTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCTTGGATGTC	734
Db	894	CAATTCAAGTGATCTCAGAGGATGCTCCTGCTTCTCAGACTGGCAGGGTCTTGGATGTC	953
Qy	735	AGTTCTGTTACAGCTAACCAAGTCATTTTGGATCTCAGAGAGAAATATTTCTAACTTAAAGCT	794
Db	954	AGTTCTGTTACAGCTAACCAAGTCATTTTGGATCTCAGAGAGAAATATTTCTAACTTAAAGCT	1013
Qy	795	CCCACAGACATCTCATAAAGCTGTGGTCAATGGAAACAATTTATGTGGGTTGTGGAGGATA	854
Db	1014	CCCACAGACATCTCATAAAGCTGTGGTCAATGGAAACAATTTATGTGGGTTGTGGAGGATA	1073
Qy	855	TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG	914
Db	1074	TATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTG	1133
Qy	915	GCTTCCACTAAACCGTTCTGTGAAACAATGTGTTTGTAGATATGGTCATTTCTTTGGCATT	974
Db	1134	GCTTCCACTAAACCGTTCTGTGAAACAATGTGTTTGTAGATATGGTCATTTCTTTGGCATT	1193
Qy	975	ATACAAGGATAAAATTTACATGTATGAGGAAAAAATTTGATTTCAACTGGGAATGTGACCAA	1034
Db	1194	ATACAAGGATAAAATTTACATGTATGAGGAAAAAATTTGATTTCAACTGGGAATGTGACCAA	1253
Qy	1035	TGAGTTGAGAGTTTTTCCACATTCATATGATGATCGGTGCTGTTGACCCCTAAGGCAA	1094
Db	1254	TGAGTTGAGAGTTTTTCCACATTCATATGATGATCGGTGCTGTTGACCCCTAAGGCAA	1313
Qy	1095	GGAGCAGTATGCAGTGTGGGCACTCTGCACACATTTGTTTACACTGAAGAATGGCCGAGT	1154
Db	1314	GGAGCAGTATGCAGTGTGGGCACTCTGCACACATTTGTTTACACTGAAGAATGGCCGAGT	1373
Qy	1155	GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAGCAATGTGCAGGA	1214
Db	1374	GGTCATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAGCAATGTGCAGGA	1433
Qy	1215	ATATGATTTGGATAAGAAACACATGGAGTATATTTACACACCCAGGGTGCCTTGTGCAAGG	1274

1434 ATATGATTGGATAAGAAACATGGAGTATATTAACACCCAGGGTGCCCTTGTGCAAGG 1493

1275 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAAGGGCCCTATACGTTTCATGTTGG 1334

1494 GGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCAAGGGCCCTATACGTTTCATGTTGG 1553

1335 CTACAAGGCTTTTCAGTGCCCAATTAAGTACCGGCTTGAGATGANTCTTACCGATATGATGT 1394

1554 CTACAAGGCTTTTCAGTGCCCAATTAAGTACCGGCTTGAGATGANTCTTACCGATATGATGT 1613

1395 GGATACCCAGATGTGGACCACTCTTAAGGACAGCCGATTTTTCGGTTACTTGCACACAGC 1454

1614 GGATACCCAGATGTGGACCACTCTTAAGGACAGCCGATTTTTCGGTTACTTGCACACAGC 1673

1455 TGTGATGTAGTGGAAACCATGCTGTTGTGGGGGAAACACACAAATGACACATCTAT 1514

1674 TGTGATGTAGTGGAAACCATGCTGTTGTGGGGGAAACACACAAATGACACATCTAT 1733

1515 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTTCAATGGCCTATGACATTCGCTGTGACCG 1574

1734 GAGCCATGGCGCCAAATGCTTCTCTTCAGATTTTCAATGGCCTATGACATTCGCTGTGACCG 1793

1575 TCTGTTCAGTGTCCCAAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCAG 1633

1794 TCTGTTCAGTGTCCCAAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATT-CAG 1852

1634 CAGTCTTACACACAGACCAATGATGTGTTTGGTGGTTTCAATAGTCTCCTCCTCAGCG 1693

1853 CAGTCTTACACACAGACCAATGATGTGTTTGGTGGTTTCAATAGTCTCCTCCTCAGCG 1912

1694 ACATCCTGGTATTCACTCTCGAACAGTGTATCGCATCGAGGTGAAGCGCGTTGTTTAG 1753

1913 ACATCCTGGTATTCACTCTCGAACAGTGTATCGCATCGAGGTGAAGCGCGTTGTTTAG 1972

1754 CAGCAGGACCTGGTATTTCGGTGTGTGAAACACAGGTTGTCTCAGTGTATCTGTTGGG 1813

1973 CAGCAGGACCTGGTATTTCGGTGTGTGAAACACAGGTTGTCTCAGTGTATCTGTTGGG 2032

1814 CGCTGGCAACTGATGACAGAGAGAAAGTTAAATCAGAATGTTTTTCCAAAGAACTC 1873

2033 CGCTGGCAACTGATGACAGAGAGAAAGTTAAATCAGAATGTTTTTCCAAAGAACTC 2092

1874 TTGACCATGACAGATGTGACCAACACACAGATTTGTACAGCTGCACAGCCAAACCAATG 1933

2093 TTGACCATGACAGATGTGACCAACACACAGATTTGTACAGCTGCACAGCCAAACCAATG 2152

1934 ACTGCCACTGGTCAATGACCATTTGTGCCAGAGAACACAGCTGCTCAGAGGGCCAGA 1993

2153 ACTGCCACTGGTCAATGACCATTTGTGCCAGAGAACACAGCTGCTCAGAGGGCCAGA 2212

1994 TCTCCATTTTAGTATGAAATGCCCCAAGGATAACCCCTATGTACTGTACTGTAAACAAGA 2053

2213 TCTCCATTTTAGTATGAAATGCCCCAAGGATAACCCCTATGTACTGTACTGTAAACAAGA 2272

2054 AGACAGGTGACAGACTGTGCCCTGGACAGAACTGCCAGTGGAGGCCCGGAAATCAGG 2113

2273 AGACAGGTGACAGACTGTGCCCTGGACAGAACTGCCAGTGGAGGCCCGGAAATCAGG 2332

2114 AGTGCAATTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAATCTCAT 2173

2333 AGTGCAATTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAATCTCAT 2392

2174 GTTTGAAAAATTACTATGCCCCAAGGAATTATGACAAATGCTAAATTTGTTCTGTAGAAACC 2233

2393 GTTTGAAAAATTACTATGCCCCAAGGAATTATGACAAATGCTAAATTTGTTCTGTAGAAACC 2452

2234 ACAATGCCCTTTTGGCTTCTCTTACAAACCAAGAGAGTAGAATTTGTCTCTTAAGCAGC 2293

2453 ACAATGCCCTTTTGGCTTCTCTTACAAACCAAGAGAGTAGAATTTGTCTCTTAAGCAGC 2512

2294 TGCGAATAATGCACTCATCTCAGAGCATGTCCAAAGCTCACTTAAACCCCATGGGTGGCC 2353



654	Db		TGTTGCCACATCAGGTTATGCCTTGCTGCAATTTTATTTAGTGATGCTGCTTTATAATTTTGAC	713
495	Qy		TGGATTTAAATTAATTAATTAACAGTTTGTATATGTGTCCAAATAAATGCTCAGGCCGACGAGA	554
714	Db		TGGATTTAAATTAATTAATTAACAGTTTGTATATGTGTCCAAATAAATGCTCAGGCCGACGAGA	773
555	Qy		GTGTAAGATCAGTAATTAAGACGCGATACGTGTTGAATGTGAATGTCTGAAATCTGAAAGG	614
774	Db		GTGTAAGATCAGTAATTAAGACGCGATACGTGTTGAATGTGAATGTCTGAAATCTGAAAGG	833
615	Qy		TGAAGCATGTGACATTCCTCACTGTACAGACAACCTGTGGTTTTCTCATCGAGGCATCTG	674
834	Db		TGAAGCATGTGACATTCCTCACTGTACAGACAACCTGTGGTTTTCTCATCGAGGCATCTG	893
675	Qy		CAATTCAAAGTCATGTCTACAGAGATGCTCCTGCTTCTCAGACTGCGACGGTCTCTGGATGTTT	734
894	Db		CAATTCAAAGTCATGTCTACAGAGATGCTCCTGCTTCTCAGACTGCGACGGTCTCTGGATGTTT	953
735	Qy		AGTTCCTGTACCAAGTCAACCAAGTCATTTTGGACTCGAGAGAAATATCTTAACCTTAAAGCT	794
954	Db		AGTTCCTGTACCAAGTCAACCAAGTCATTTTGGACTCGAGAGAAATATCTTAACCTTAAAGCT	1013
795	Qy		CCCAGAGCATCTCATAAAGCTGTGGTCAATTGAAACAATTAATGTGGGTCTCTCGAGGATA	854
1014	Db		CCCAGAGCATCTCATAAAGCTGTGGTCAATTGAAACAATTAATGTGGGTCTCTCGAGGATA	1073
855	Qy		TATGTTCAACCACTCAGATTAACATGTTTTAGCGTATGACCTTGCTTCTTAGGAGTG	914
1074	Db		TATGTTCAACCACTCAGATTAACATGTTTTAGCGTATGACCTTGCTTCTTAGGAGTG	1133
915	Qy		GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTGTAGATATGTCATCTTTTGCATT	974
1134	Db		GCTTCCACTAAACCGTTCTGTGAACAATGTGGTTGTGTAGATATGTCATCTTTTGCATT	1193
975	Qy		ATACAAGGATAAAATTTACATGTATGGAGAAAATTTGATTCAACTGGGAATGTGACCAA	1034
1194	Db		ATACAAGGATAAAATTTACATGTATGGAGAAAATTTGATTCAACTGGGAATGTGACCAA	1253
1035	Qy		TGAGTTGAGAGTTTTTACAAATTCATAATGAGTCATGGGTGTTCTTGACCCCTTAAGCCAA	1094
1254	Db		TGAGTTGAGAGTTTTTACAAATTCATAATGAGTCATGGGTGTTCTTGACCCCTTAAGCCAA	1313
1095	Qy		GGAGCAGTATGCAAGTGTGTGGGCACCTGACACACATTTGTACACTGAAGAATGGCCGAGT	1154
1314	Db		GGAGCAGTATGCAAGTGTGTGGGCACCTGACACACATTTGTACACTGAAGAATGGCCGAGT	1373
1155	Qy		GGTCATGCTGGTCAATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGAGGA	1214
1374	Db		GGTCATGCTGGTCAATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGTGAGGA	1433
1215	Qy		ATATGATTTGGATAAGAAACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG	1274
1434	Db		ATATGATTTGGATAAGAAACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG	1493
1275	Qy		GGGTTACGGCCATAGCAGTGTTTACGACCAATAGAACCCAGGGCCCTATACGTTCAATGGTGG	1334
1494	Db		GGGTTACGGCCATAGCAGTGTTTACGACCAATAGAACCCAGGGCCCTATACGTTCAATGGTGG	1553
1335	Qy		CTACAAGCTTTTCAGTGGCCAAATAAGTACCGGCTTGAGATGATCTCTACCCGATATGATGT	1394
1554	Db		CTACAAGCTTTTCAGTGGCCAAATAAGTACCGGCTTGAGATGATCTCTACCCGATATGATGT	1613
1395	Qy		GGATACCCAGATGTGGACCAATCTTAAGGACAGCCGATTTTTCCGTTACTTTGCACACAGC	1454
1614	Db		GGATACCCAGATGTGGACCAATCTTAAGGACAGCCGATTTTTCCGTTACTTTGCACACAGC	1673
1455	Qy		TGTGATATGAGTGGACCAATGCTGGTGTGGGGGAAACACACAAATGACACATCTAT	1514
1674	Db		TGTGATATGAGTGGACCAATGCTGGTGTGGGGGAAACACACAAATGACACATCTAT	1733
1515	Qy		GAGCCATGGCGCCAAATGCTCTCTTTTCAGATTTTCTAGGCCCTATGACATTTGCTGTGACCG	1574

1734	Db	GAGCCATGGCGCCAAANTGCTTCTCTTCACAGATTTCATGGCCTATGACATTTGCTCCTGTGACCG	1799
1575	Qy	CTGGTCAGTGTCTCCAGGACCTTGAT-TCCACCAATGATGTCAACAGAGATTTGGCCATTTCCAG	1633
1794	Db	CTGGTCAGTGTCTCCAGGACCTTGATCTCCACCAATGATGTCAACAGAGATTTGGCCATTTCCAG	1852
1634	Qy	CAGTCTTTACACACAGCAGCAGATGTATGTGTTCCGTGGTGTTCATAGTCTCTCTCTCTCAGCG	1693
1853	Db	CAGTCTTTACACACAGCAGCAGATGTATGTGTTCCGTGGTGTTCATAGTCTCTCTCTCTCAGCG	1912
1694	Qy	ACATCTCTGTATTTCACCTCGGAACACAGTGTGATCGGCATCGGAGTGAAGCGCTTGTTTAG	1753
1913	Db	ACATCTCTGTATTTCACCTCGGAACACAGTGTGATCGGCATCGGAGTGAAGCGCTTGTTTAG	1972
1754	Qy	CAGCAGGACCTGTGATTTCCGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCTGTGG	1813
1973	Db	CAGCAGGACCTGTGATTTCCGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCTGTGG	2032
1814	Qy	CGCTGCGCAACTGATGAACAAGAAAGAGTTAAATTCAGAAATGTTTTTCACAAAGAACTC	1873
2033	Db	CGCTGCGCAACTGATGAACAAGAAAGAGTTAAATTCAGAAATGTTTTTCACAAAGAACTC	2092
1874	Qy	TTGACCATGACAGATGTGACAGCACACAGATTTGTTTACAGCTGACACAGCCAAACACCAATG	1933
2093	Db	TTGACCATGACAGATGTGACAGCACACAGATTTGTTTACAGCTGACACAGCCAAACACCAATG	2152
1934	Qy	ACTGCCACTGGTGCATGACCATTGTTGTCGCCAGGAACACACAGCTGCTCAGAGGCCAGA	1993
2153	Db	ACTGCCACTGGTGCATGACCATTGTTGTCGCCAGGAACACACAGCTGCTCAGAGGCCAGA	2212
1994	Qy	TCTCCATTTTTAGTATGAGAAATGGCCCCAAGGATAACCTATGTACTACTGTAAACAAGA	2053
2213	Db	TCTCCATTTTTAGTATGAGAAATGGCCCCAAGGATAACCTATGTACTACTGTAAACAAGA	2272
2054	Qy	AGACAGCTGCGAGGAGCTGTGCCCTGGACACAGAACTGCCAGTGGGAGCCCGCGAATCAGG	2113
2273	Db	AGACAGCTGCGAGGAGCTGTGCCCTGGACACAGAACTGCCAGTGGGAGCCCGCGAATCAGG	2332
2114	Qy	AGTGCATTCGCCCTGCCCCGAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAACCTCAT	2173
2333	Db	AGTGCATTCGCCCTGCCCCGAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAACCTCAT	2392
2174	Qy	GTTTGMAAATTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACC	2233
2393	Db	GTTTGMAAATTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACC	2452
2234	Qy	ACAATGCCCTTTTGGCTTCTCTTACAAACCAGAGAGAGTGAATTTGTCTTTAAGCAGC	2293
2453	Db	ACAATGCCCTTTTGGCTTCTCTTACAAACCAGAGAGAGTGAATTTGTCTTTAAGCAGC	2512
2294	Qy	TGCGAATAATGCAGTCACTCTCAGAGCATGTCCAAGCTCACCTTAAACCCCATGGTGGGCC	2353
2513	Db	TGCGAATAATGCAGTCACTCTCAGAGCATGTCCAAGCTCACCTTAAACCCCATGGTGGGCC	2572
2354	Qy	TTCCGAAGATCAATGTGTCTACTGGTGTCTGGAGAGATGTGCCCATTTTACAAATAGTT	2413
2573	Db	TTCCGAAGATCAATGTGTCTACTGGTGTCTGGAGAGATGTGCCCATTTTACAAATAGTT	2632
2414	Qy	TACTACAGTGGATGCCCCTCTGAGCCCAAGTGTGATGCTGGATTTCTGTGGAAATTTTATCAGAAC	2473
2633	Db	TACTACAGTGGATGCCCCTCTGAGCCCAAGTGTGATGCTGGATTTCTGTGGAAATTTTATCAGAAC	2692
2474	Qy	CCAGTACTCGGGGACATGAAGGCTGCAACTGTGATCAACCCACTCAATGCTAGTGTCTGTG	2533
2693	Db	CCAGTACTCGGGGACATGAAGGCTGCAACCTGTGATCAACCCACTCAATGCTAGTGTCTGTG	2752
2534	Qy	AAAGGCTCGCAAAACCAACAGTGTGAACAGTGGCGGACACCATGTGCTTGAGGACACAGCAT	2593
2753	Db	AAAGGCTCGCAAAACCAACAGTGTGAACAGTGGCGGACACCATGTGCTTGAGGACACAGCAT	2812
2594	Qy	GTGGAGATTTGCACAGCGGACGCTCTGAGTGTGATGTGGTGACAGCAACATGAAGCAGTGTG	2653
2813	Db	GTGGAGATTTGCACAGCGGACGCTCTGAGTGTGATGTGGTGACAGCAACATGAAGCAGTGTG	2872

2654 TGGACTCCAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGA 2713  
2873 TGGACTCCAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGA 2932  
2714 TGGACACTGCCCCCTGAAAAATGTTTCAGGCTACTGTACTCTAGTCTACTGTTGGAGC 2773  
2933 TGGACACTGCCCCCTGAAAAATGTTTCAGGCTACTGTACTCTAGTCTACTGTTGGAGC 2992  
2774 AACAGGCTGTGGTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGCATAGAGG 2833  
2993 AACAGGCTGTGGTGTGTACTGATCCAGCAATATCTGGCAAGGGAATGCATAGAGG 3052  
2834 GTTCCCTATAAGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAC 2893  
3053 GTTCCCTATAAGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAC 3112  
2894 AGCCCCCTGCTCAATTCAGCATGTGTCTAGAGACACAGCATCAATCTGGTCTTTCAATC 2953  
3113 AGCCCCCTGCTCAATTCAGCATGTGTCTAGAGACACAGCATCAATCTGGTCTTTCAATC 3172  
2954 ACTGTCCAGCTTCCATGCAAGCGCCACAGTAAATGCAATCAATCAGACATCTGTGAGA 3013  
3173 ACTGTCCAGCTTCCATGCAAGCGCCACAGTAAATGCAATCAATCAGACATCTGTGAGA 3232  
3014 AGTGTGAGAACCTGACACAGCAAGCACTGCGAGACCTGTCATATCTGGCTTCTACGGTG 3073  
3233 AGTGTGAGAACCTGACACAGCAAGCACTGCGAGACCTGTCATATCTGGCTTCTACGGTG 3292  
3074 ATCCACCAATGAGGGAATGTGAGCATGCAAGTGAATGCAATCAATCAGACATCTGTGAGA 3133  
3293 ATCCACCAATGAGGGAATGTGAGCATGCAAGTGAATGCAATCAATCAGACATCTGTGAGA 3352  
3134 ACACCAACCGGCAAGTGTCTGTCACCAACCAAGCGCGTCAAGGGGAGCGAGTGCAGC 3193  
3353 ACACCAACCGGCAAGTGTCTGTCACCAACCAAGCGCGTCAAGGGGAGCGAGTGCAGC 3412  
3194 TATGTAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAAATGTTATATATCTC 3253  
3413 TATGTAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAAATGTTATATATCTC 3472  
3254 TTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAAAGATGATCGCTATTACACAG 3313  
3473 TTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAAAGATGATCGCTATTACACAG 3532  
3314 CTATCAATTTTGTGGTACTCTGACGAAACAAACAGGGATTTGGACATGTTTCATCAATG 3373  
3533 CTATCAATTTTGTGGTACTCTGACGAAACAAACAGGGATTTGGACATGTTTCATCAATG 3592  
3374 CCTCCAGAAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433  
3593 CCTCCAGAAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652  
3434 CTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAGATAGTTTCT 3493  
3653 CTGGAGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAGATAGTTTCT 3712  
3494 CTAATGAGAGTTTGTATTTTCGAAACCAACCAATATCACTTTCTTTGTTATGTGAGTA 3553  
3713 CTAATGAGAGTTTGTATTTTCGAAACCAACCAATATCACTTTCTTTGTTATGTGAGTA 3772  
3554 ATTTTCACCTGGCCCATCAAAATTCAGGT 3581  
3773 ATTTTCACCTGGCCCATCAAAATTCAGAT 3800

## RESULT 3

US-09-893-238-1

; Sequence 1, Application US/09893238

; Patent No. US2020150973A1

; GENERAL INFORMATION:

; APPLICANT: Moore, K.

; APPLICANT: Nagle, D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8827  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; US-09-893-238-1

Query Match 79.1%; Score 2845.2; DB 9; Length 8827;  
Best Local Similarity 90.3%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 328; Indels 2; Gaps 2;  
QY 189 ACTAAGTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAATTAATAATACAAACGAA 248  
Db 677 ACTAAGTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAATTAATAATACAAACGAA 736  
QY 249 GTGACGCTGGCTCATTTGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATCATTT 308  
Db 737 GTGACATGGCTCATTTGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATCATTT 796  
QY 309 TGCTACAGAGTGTAGTTGGGACCAATTAATGTTTATGATGGGACTCAATTTATGACAC 368  
Db 797 TGCTACAGAGTGTAGTTGGGACCAATTAATGTTTATGATGGGACTCAATTTATGACAC 856  
QY 369 GCTAGTCTGCTCATTTAGTGGCTCATTTGCTGAGAGATGCAATGAGACTGTGCTCC 428  
Db 857 TCTGATCTGCTCATTTAGTGGCTCATTTGCTGAGAGATGCAATGAGACTGTGCTCC 916  
QY 429 TGAGTGTGTTGCAATCAGGTTATGCTTCTGCAATTTTATGATGCTGCTTATAA 488  
Db 917 TGAGTGTGTTGCAATCAGGTTATGCTTCTGCAATTTTATGATGCTGCTTATAA 976  
QY 489 TTTGACTGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 548  
Db 977 TCTGATGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1036  
QY 549 AGGAGAGTGAAGATCAGTAAATAGCAGCGATACCTTTGAATGTAATGTTCTGAAAAC 608  
Db 1037 AGGAGAGTGAAGATCAGTAAATAGCAGCGATACCTTTGAATGTAATGTTCTGAAAAC 1096  
QY 609 GAAAGTGAAGCATGTGACATTTCTCAGTGTACAGACAACTGTGTTTCTCATCGAGG 668  
Db 1097 GAAAGTGAAGCATGTGACATTTCTCAGTGTACAGACAACTGTGTTTCTCATCGAGG 1156  
QY 669 CATCTGCAATTCAGAGTGTGAGAGGATGCTTCTGCTTCTCAGACTGGCAGGCTCTGG 728  
Db 1157 CATCTGCAATTCAGAGTGTGAGAGGATGCTTCTGCTTCTCAGACTGGCAGGCTCTGG 1216  
QY 729 ATGTTTCACTTCTGTAACAGCTTAACAGTCAATTTTGGACTCGAGAGGAATATCTTAAT 788  
Db 1217 ATGTTTCACTTCTGTAACAGCTTAACAGTCAATTTTGGACTCGAGAGGAATATCTTAAT 1276  
QY 789 AAAGTCTCCAGAGCATCTCAATTAAGCTGTGCTCAATGGAACATTAATGTTGGTTTGG 848  
Db 1277 AAAGTCTCCAGAGCATCTCAATTAAGCTGTGCTCAATGGAACATTAATGTTGGTTTGG 1336  
QY 849 AGGATATATGTTCAACCACTCAGATTAATAACATGTTCTAGCGTATGACCTTGTCTTAG 908  
Db 1337 CGGATATATGTTCAACCACTCAGATTAATAACATGTTCTAGCGTATGACCTTGTCTTAG 1396  
QY 909 GGAGTGGCTTCCACTAAACCGTCTGTGAACATGTTGTTGTAGATATGTCATCTTT 968



Db 1397 GGAATGGCTTCCACTAAACCAATTCCTGTGAACAGTGTGGTTGTGAAGATATAGGTCAATCTTT 1456  
Qy 969 GGCAATTATACAAGGATAAAATTTACATGTATGAGAGAAAATTTGATTTCAACTCGGAATGT 1028  
Db 1457 GGCAATTATACAAGGATAAAATTTACATGTATGAGAGAAAATTTGATTTCAACAGGAACTGT 1516  
Qy 1029 GACCAATGAGTTGAGAGTTTTCATCATTAATGAGTCAATGGGTGTGTGACCCCTAA 1088  
Db 1517 GACCAATGAGTTGAGAGTTTTCATCATTAATGAGTCAATGGGTGTGTGACCCCTAA 1576  
Qy 1089 GGCAAGGAGCAGTATCGAGTGTGGGCACTCTGCACATTTGTACACTGAAGAATGG 1148  
Db 1577 AGCTTAGGATCAGTATGAGAGTGTGGACACTCAGCAATTTGTACACTGGCATCTGG 1636  
Qy 1149 CCGAGTGGTCATGCTGTCTATTTGTGTCACTGCGCTCTCTATGGATATATAAGCAATGT 1208  
Db 1637 CCGTGTGGTCATGCTGTCTATTTGGTCTATTTGGCTCATTTGCCACTCTATGGATATATAAGCGTTGT 1696  
Qy 1209 GCAGGATATGATTTGGATGAGACATCGAGTATATTTACACACCCAGGCTGCCCTGT 1268  
Db 1697 GCAGGATATGACTTTGGAAGAAACACATGAGTATATTTACATCTCAGGCTGCTTTGT 1756  
Qy 1269 GCAAGGGGGTTACCGGCATAGCAGTGTTTACGACCATAGGACAGGCGCCCTATACGTTCA 1328  
Db 1757 GCAAGGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816  
Qy 1329 TGGTGGCTACAGGCTTTCACTGSCAATAGTACCGGCTTGCAGATGATCTCTACCGATA 1388  
Db 1817 TGGTGGCTACAGGCTTTCACTGSCAATAGTACCGGCTTGCAGATGATCTCTACCGATA 1876  
Qy 1389 TGATGTGATACCCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTGCA 1448  
Db 1877 CGATGTGATCTCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTGCA 1936  
Qy 1449 CACAGCTGTGATAGTGTGAACCAATGCTGGTGTGTTGGGGGAAACACACAAATGACAC 1508  
Db 1937 TACAGCTGTGATAGTGTGAACCAATGCTGGTGTGTTGGGGGAAACACACAAATGACAC 1996  
Qy 1509 ATCTATGAGCCATGGCGCCAAATGCTTCTCTCAGATTTTCATGCTCTATGACATTCCTCG 1568  
Db 1997 TTCCATGAGCCACGGTCCAAATGCTTCTCTCAGATTTTCATGCTCTATGACATTCCTCG 2056  
Qy 1569 TGACCGCTGTGCTGCTTCCAGACCTGA - TTCCACCAATGATGTCAACAGATTTGGCCA 1627  
Db 2057 TGACCGATGTGCTGCTTCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCA 2116  
Qy 1628 TTCCAGCAGCTTTACACACAGCCATGATGTGTGCTGGTGTTCATAGTCTCTCTCC 1687  
Db 2117 TT - CAGCAGCTTTGTACACAGCACCATGATGTGTGCTGGCCTTCAACAGGCTCTCTCC 2175  
Qy 1688 TCAGCAGCATCTCTGCTATTCACCTCGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTT 1747  
Db 2176 TCAGTGAAGCTTTGCTTTTACCTCGAGCAGTGGCATGACACCCGAGTGAAGCTGCTT 2235  
Qy 1748 GTTTAGCAGCAGGACCTGTTATTTCCGTTGTGTGGAACACAGGCTGTCTCAGTGTATCT 1807  
Db 2236 GTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACAGTGTCTCTGATGTACCT 2295  
Qy 1808 CGTGGGGCTGGCAACTGATGAACACAGAAAAGTTTAAATCAGAAATGTTTTTCCAAA 1867  
Db 2296 CCTGGAGTTGGCACTGAGAACACAGCAAGTAAATCAGAGTGTTTTTTCTAAA 2355  
Qy 1868 GAATCTTTGACATGACAGATGTGACAGCAGACAGATTTGTTACAGTGTGACAGCAACA 1927  
Db 2356 GAACCTTTGACATGACAGATGTGACAGCAGACAGATTTGTTACAGTGTGACAGCAACA 2415  
Qy 1928 CCAATGATCTGCACTGGTGCAATGAACATTTGTGCTCCAGGAACCCAGCTGTCTCAGAAG 1987  
Db 2416 CCAATGATCTGCACTGGTGCAATGATCACTGTCTCTGTGTGTAACCAAGCTGTGACAGAG 2475  
Qy 1988 GCCAGATCTCCATTTTAGGTATGAGAAATTTGCCCCAAAGGATATCCCTATGTACTCTGTA 2047  
Db 2476 GCCAGATCTCCATTTAGGTATGAGAAATTTGCCCCAAAGGATATCCCTATGTACTCTGTA 2535

Qy 2048 ACAAGAAGACAGCTGTCAGGAGCTGTGCCCTGGACCAAGAACTGCAGTGGAGAGCCCCGA 2107  
Db 2536 ATAAGAAAACAGCTGTCAGGAGCTGTGCCCTAGACCAAGAACTGCAGTGGAGAGCCCCGA 2595  
Qy 2108 ATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGGCAATTTGGCTGCAATTTGTTGGAA 2167  
Db 2596 ATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGGCAATTTGGCTGCAATTTGTTGGAA 2655  
Qy 2168 ACTCATGTTTCAAAATTTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTA 2227  
Db 2656 ACTCGTGTCTGAAATCACTACTGCTAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTA 2715  
Qy 2228 GGAACCAAAATGCTGCTTTGGCTTCTCTTACAAACCAAGGAGAAATTTGTTCTGTA 2287  
Db 2716 GGAACCAAAATGCTGCTTTGGCTTCTCTTACAAACCAAGGAGAAATTTGTTCTGTA 2775  
Qy 2288 AGCAGCTCGGAATTAATGCTGCTCAGAGCATCTCCAAGCTCACCTTAAACCCCATGGG 2347  
Db 2776 AGCAGCTTCGATTAATGCAATCACTCAAGTATGTCCAAGCTCACTCTGACCTCATGGG 2835  
Qy 2348 TCGGCTTTCCGAAAGATCAATGTGCTTACTTGGTGTGGAAGATATGTCCCATTTACAA 2407  
Db 2836 TTGTGTTTCCGAAAGATCAATGTGCTTACTTGGTGTGGAAGATATGTCTCCATTTCAA 2895  
Qy 2408 ATAGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTGTGGAATTTCTGTGGAATTTTAT 2467  
Db 2896 ATAGTTTCTGAGTGGATGCCATCTGAGCCCAAGTGTGGAATTTCTGTGGAATTTTAT 2955  
Qy 2468 CAGAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTG 2527  
Db 2956 CAGAGCTTAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCCTCAATGGTAGTG 3015  
Qy 2528 TCTGTGAAGCCCTGCAAAACCAACAGTGTGAGAGTGCAGACACCATGTGCTTTGAGGA 2587  
Db 3016 TCTGTGAAGCCCTGCAAAACCAACAGTGTGAGAGTGCAGACACCATGTGCTTTGAGGA 3075  
Qy 2588 CAGCATGTGGAGATTTGCAACCAAGCCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGC 2647  
Db 3076 CAGCGTGTGGGAGTGCACATAGCAGAGCTCGGAGTGCATGTGGTGCAGCAACATGAAGC 3135  
Qy 2648 AGTGTGTGACTCMAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTGGAATGGT 2707  
Db 3136 AGTGTGTGACTCMAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTGGAATGGT 3195  
Qy 2708 ATACATGAGCAGCTGCCCCCTGAAAATTTGTTCAAGCTACTGTACCTGTAGTCAATGTCT 2767  
Db 3196 ATACATGAGCAGCTGCCCCCTGAAAATTTGCTCTGCTACTGTACCTGCAGCCATGTCT 3255  
Qy 2768 TGGAGCAACCAAGCTGTGGTGTGATGATCCAGCAATATCTGGAAGGAAATGCA 2827  
Db 3256 TGGAGCAGCAGGCTGTGGTGTGATGATCCATGATCTGGAAGGAAATGCA 3315  
Qy 2828 TAGAGGGTTCTATAAAGCAGCAGTGAAGTGCCTTGCAGAGCCCTTACAGGAAATTTCT 2887  
Db 3316 TTGAGGGCAGCTATAAAGCACTGTGAAGTGCCTTACAGCCCTCTGCAGGAAATGTCT 3375  
Qy 2888 ATCCACAGCCCTGCTCAATTTCCAGCATGTGTGTAGAGGACAGCAGATACACTGTGCTT 2947  
Db 3376 ATCCACAGCCCTTCTGAATCTCAGCATGTGTGTAGAGGACAGCAGATACACTGTGCTT 3435  
Qy 2948 TCATTTCACTGTCCAGCTTGCCTAATGCAACGGCCACAGTAAATGATCAATCAGAGCATCT 3007  
Db 3436 TCATTTCACTGTCCAGCTTGCCTAATGCAACGGCCACAGCAGCAATGATCAACAGATATCT 3495  
Qy 3008 GTGAGAGTGTGAGAACTGACCAACAGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3067  
Db 3496 GTGAGAGTGTGAGAACTGACCAACAGCAAGCACTGCGAGACCTGCATATCTGGCTTCT 3555  
Qy 3068 ACGGTGATCCCAACCAATGAGGAGAAATGTGAGCCATGCAATGCAATGGGACCGCTCTC 3127  
Db 3556 ATGGTGACCCGACTAATGGAGGCAATGTGAGCCATGCAATGCAATGGGACCGCATCAC 3615

```
QY 3128 TGTGCAACACCAACGCGGAGTGTCTCTGACACCAACGAGCGGTCAAGGGGAGGAGT 3187
Db 3616 TGTGCAACACCAACGCGGAGTGTCTCTGACACCAACGAGGTGTCAGAGGGGAGGAGT 3675
QY 3188 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATT 3247
Db 3676 GCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGCTACT 3735
QY 3248 ATACTCTTCTTATGACTATCAGTTCACTTATGCTTATCTTCCAGGAGAGATGATCGCTATT 3307
Db 3736 ATACCCCTTCTCAATGACTATCAGTTCACTTATGCTTATCTTCCAGGAGAGATGATCGCTACT 3795
QY 3308 ACACAGCTATCAATTTTGGCTACTCTCAGCAACAAACAGGAGTTTGGACATGTTCA 3367
Db 3796 ACACAGCTATCAATTTTGGCTACTCTCAGCAACAAACAGGAGTTTGGACATGTTCA 3855
QY 3368 TCAATGCTTCAAGAAATTTCAACTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAA 3427
Db 3856 TCAATGCTTCAAGAAATTTCAACTCAACATCACTGGGCTGCCAGTTTCCAGCGGAA 3915
QY 3428 CCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATA 3487
Db 3916 CCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATA 3975
QY 3488 GTTCTCTCAATGAGAAATTTGATTTTCCGAACCAACCAATATCACTTTCTTTGTTTATG 3547
Db 3976 GTTCTCTCAATGAGAAATTTGATTTTCCGAACCAACCAATATCACTTTCTTTGTTTATG 4035
QY 3548 TCAGTAATTTCACTGGCCCATCAAAATTCAGT 3581
Db 4036 TCAGTAATTTCACTGGCCCATCAAAATTCAGAT 4069
```

## RESULT 4

```
US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; ; FILE REFERENCE: 7853-237
; ; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; ; CURRENT APPLICATION NUMBER: US/09/893,238
; ; CURRENT FILING DATE: 2001-06-27
; ; PRIOR APPLICATION NUMBER: 09/245,041
; ; PRIOR FILING DATE: 1999-02-05
; ; PRIOR APPLICATION NUMBER: 60/093,630
; ; PRIOR FILING DATE: 1998-07-21
; ; PRIOR APPLICATION NUMBER: 60/104,978
; ; PRIOR FILING DATE: 1998-10-20
; ; NUMBER OF SEQ ID NOS: 129
; ; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18
```

Query Match 56.4%; Score 2028.8; DB 9; Length 2625;  
Best Local Similarity 97.0%; Pred. No. 0;  
Matches 2088; Conservative 0; Mismatches 62; Indels 2; Gaps 2;

```
QY 15 GCGCGCAACTGAGGCAAGGTGAGGAGGAGCGCGGCGACGCGCGCTCGCGGCGAG 74
Db 234 GGAGGCGAGCGCGCTGCGGTGGCGGCGGTGTCTCGGCTCGCGCGCGAGCGGCA 293
QY 75 GAGCGGCGCGCGCTGCTCAACGCGGTGCTCAACCTGGACCGCGCGAGTGCCT 134
Db 294 GGAATGTGACCGCGGTGTGTCAACGCGGTGCTGCAACCTGGACCGCGCGAGTGCCT 353
QY 135 CTGCCCCCGCGGTGGGTGGCGAGCAATGCCAGCACTCGCGGGGCGGCTTCAGACTAAC 194
```

```
Db 354 CTGCCCCCGCGGTGGGTGGCGAGCAATGCCAGCACTCGCGGGGCGGCTTCAGACTAAC 413
QY 195 TGGATCTTCTGGGTTTGTGACAGATGGACCTCGGAAATTAATAAACAAGAGTGCAC 254
Db 414 TGGATCTTCTGGGTTTGTGACAGATGGACCTCGGAAATTAATAAACAAGAGTGCAC 473
QY 255 GTGCTCAATTCAGAGCAGCCAAATAGAATAATCAGACTTCGTTTCAATCATTTTGTCTAC 314
Db 474 GTGCTCAATTCAGAGCAGCCAAATAGAATAATCAGACTTCGTTTCAATCATTTTGTCTAC 533
QY 315 AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGAGCTCAATTTATGACCCTAGT 374
Db 534 AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGAGCTCAATTTATGACCCTAGT 593
QY 375 TGCTGCATTTAGTGGCCCTCATTTCTCTGAGAGAGTGGCAATCAGACTGTCCTGAGGT 434
Db 594 TGCTGCATTTAGTGGCCCTCATTTCTCTGAGAGAGTGGCAATCAGACTGTCCTGAGGT 653
QY 435 TGTGGCCACATCAGGTTATGCTTGTCTGCATTTTTTTTAGTGATGCTGTTATATATTGAC 494
Db 654 TGTGGCCACATCAGGTTATGCTTGTCTGCATTTTTTTTAGTGATGCTGTTATATATTGAC 713
QY 495 TGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 554
Db 714 TGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 773
QY 555 GTGTAAGATCAGTAATAGCAGCGATACCTGTTGAATGTGAATGTTCTGAAAACCTGAAAGG 614
Db 774 GTGTAAGATCAGTAATAGCAGCGATACCTGTTGAATGTGAATGTTCTGAAAACCTGAAAGG 833
QY 615 TGAAGATGTGACATTTCTCACTGTACAGACAACTGTGGTTTTTCTCATCAGAGCACTGTG 674
Db 834 TGAAGATGTGACATTTCTCACTGTACAGACAACTGTGGTTTTTCTCATCAGAGCACTGTG 893
QY 675 CAATTCAGTGTGTCAGAGGATGCTCTGCTCTCAGACTGCGAGGCTGCTGATGTC 734
Db 894 CAATTCAGTGTGTCAGAGGATGCTCTGCTCTCAGACTGCGAGGCTGCTGATGTC 953
QY 735 AGTTCCTGTACCACTAACCACTATTTTGGACTCGAGAGGAATATTCTAACTTTAAAGCT 794
Db 954 AGTTCCTGTACCACTAACCACTATTTTGGACTCGAGAGGAATATTCTAACTTTAAAGCT 1013
QY 795 CCCAGAGCTCTCATAAAGCTGTGGTCAATGGAACATTAATGTTGGGTGTTGGAGGATA 854
Db 1014 CCCAGAGCTCTCATAAAGCTGTGGTCAATGGAACATTAATGTTGGGTGTTGGAGGATA 1073
QY 855 TATGTTCAACCACTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 914
Db 1074 TATGTTCAACCACTCAGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1133
QY 915 GCTTCCACTAAACCGTCTGTGAACATGTTGGTTGTTAGATATGTCATTTCTTTGGCAATT 974
Db 1134 GCTTCCACTAAACCGTCTGTGAACATGTTGGTTGTTAGATATGTCATTTCTTTGGCAATT 1193
QY 975 ATACAAGGATAAAATTTACATGTTGGAGGAAATTAATTAATTAATTAATTAATTAATTAATTA 1034
Db 1194 ATACAAGGATAAAATTTACATGTTGGAGGAAATTAATTAATTAATTAATTAATTAATTAATTA 1253
QY 1035 TGAGTTGAGAGTCTTTTTCATTTTCTGCTCTCTATGATGATATTAAGCAATTTGCGAGGA 1094
Db 1254 TGAGTTGAGAGTCTTTTTCATTTTCTGCTCTCTATGATGATATTAAGCAATTTGCGAGGA 1313
QY 1095 GGAGGAGTATCAGTGGTTGGGCACTCTGCAACATTTGTTACATGAGATGCGCGAGT 1154
Db 1314 GGAGGAGTATCAGTGGTTGGGCACTCTGCAACATTTGTTACATGAGATGCGCGAGT 1373
QY 1155 GGTCAATGCTGCTCATCTTTTGGTCACTGCTCTCTATGATGATATTAAGCAATTTGCGAGGA 1214
Db 1374 GGTCAATGCTGCTCATCTTTTGGTCACTGCTCTCTATGATGATATTAAGCAATTTGCGAGGA 1433
QY 1215 ATATGATTTGATTAAGAACATGAGGATATTAATTAACACCCAGGCTGCTTTGCGAAGG 1274
Db 1434 ATATGATTTGATTAAGAACATGAGGATATTAATTAACACCCAGGCTGCTTTGCGAAGG 1493
```

1275 GGGTTACGGCCATAGCAGTGTTCACGACCATAGACACAGGCGCCCTATACGTTTCATCGTGG 1334  
Db GGGTTACGGCCATAGCAGTGTTCACGACCATAGACACAGGCGCCCTATACGTTTCATCGTGG 1553  
QY CTACAGGCTTTCAGTCCCAATAGTACCGGCTGAGATGATCTCTACCGATATGATGT 1394  
Db CTACAGGCTTTCAGTCCCAATAGTACCGGCTGAGATGATCTCTACCGATATGATGT 1613  
QY GGATACCCAGATGAGACCATCTTAAAGGACAGCCGATTTTCGGTTACTTGCACACAGC 1454  
Db GGATACCCAGATGAGACCATCTTAAAGGACAGCCGATTTTCGGTTACTTGCACACAGC 1673  
QY TGTGATAGTACGAGACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTAT 1514  
Db TGTGATAGTACGAGACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTAT 1733  
QY GAGCCATGGGCGCCAAATGCTTCTTCAGATTTTCATGGCCCTATGACATTCGCTGTGACCG 1574  
Db GAGCCATGGGCGCCAAATGCTTCTTCAGATTTTCATGGCCCTATGACATTCGCTGTGACCG 1793  
QY CTGTCAGTGTTCCTCCAGACCTGAT-TCCACCATGATGTCAAAGATTTGGCCATTCGAG 1633  
Db CTGTCAGTGTTCCTCCAGACCTGAT-TCCACCATGATGTCAAAGATTTGGCCATTCGAG 1852  
QY CAGTCTTACACACAGCAGCAGCATGATGTGTTGGTGTGTTCAATAGTCTCTCTCCTCAGCG 1693  
Db CAGTCTTACACACAGCAGCAGCATGATGTGTTGGTGTGTTCAATAGTCTCTCTCCTCAGCG 1912  
QY ACATCTCTGTTATTCACCTCGGAAACAGTGTGATGCGATCGAGTGAAGCGCTGTGTTAG 1753  
Db ACATCTCTGTTATTCACCTCGGAAACAGTGTGATGCGATCGAGTGAAGCGCTGTGTTAG 1972  
QY CAGCAGACCTGTTATTCGGTGTGTTGGAAACAGGCTCGTCTCAGTGTATCTCGTGGG 1813  
Db CAGCAGACCTGTTATTCGGTGTGTTGGAAACAGGCTCGTCTCAGTGTATCTCGTGGG 2032  
QY CGCTGGCACTGATGAACAGAAAGAAAGTTAAATCAGAAATGTTTTCCAAAGAACTC 1873  
Db CGCTGGCACTGATGAACAGAAAGAAAGTTAAATCAGAAATGTTTTCCAAAGAACTC 2092  
QY TTGACCATGACAGATGACAGCAGCAGATGTTACAGCTGACAGCAGCAGCAGCAATG 1933  
Db TTGACCATGACAGATGACAGCAGCAGATGTTACAGCTGACAGCAGCAGCAGCAATG 2152  
QY ACTGCCACTGTGCAATGACCATTTGTGTCCTCCAGGAAACACAGCTGCTCAGAAAGGCCAG 1993  
Db ACTGCCACTGTGCAATGACCATTTGTGTCCTCCAGGAAACACAGCTGCTCAGAAAGGCCAG 2212  
QY TCTCCATTTTAGTATGAGATTTGCCCAAGGATTAACCTATGTAATGTAATGTAACAG 2053  
Db TCTCCATTTTAGTATGAGATTTGCCCAAGGATTAACCTATGTAATGTAATGTAACAG 2272  
QY AGACCACTGACAGAGCTGTGCTGACAGCAAGCTGCGAGTGGAGCCCGGAAATCAGG 2113  
Db AGACCACTGACAGAGCTGTGCTGACAGCAAGCTGCGAGTGGAGCCCGGAAATCAGG 2332  
QY AGTGCATTCGCTCCCGAAATATCTGTGGCAATTCGCTGGCAATTTGGTTGG 2165  
Db AGTGCATTCGCTCCCGGAGGCTTTCAGGGTCTATCTGTTGTTGTTGGG 2384

## RESULT 5

US-09-893-238-8  
; Sequence 8, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238

; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-8

Query Match 39.7%; Score 1428.8; DB 9; Length 2419;  
Best Local Similarity 89.4%; Pred. No. 0;  
Matches 1560; Conservative 0; Mismatches 182; Indels 2; Gaps 2;

QY 189 ACTAACTGATCTTCTGGTTTGTGACAGATGGACCTGGAAATATATAAATACAAACGAA 248  
Db 677 ACTAACTGATCTTCTGGTTTGTGACAGATGGACCTGGAAATATATAAATACAAACGAA 736  
QY 249 GTGCACGTGGCTCAATGAAGACAGCCAAATAGAAATATAGACTTCTGTTTCAATCATTT 308  
Db 737 GTGCACATGGCTCAATGAAGACAGCCAAATAGAAATATAGACTTCTGTTTCAACCATTT 796  
QY 309 TGCTACAGATGTAGTGGACCAATTTATATGTTTATGATGGGACCTCAATTTATGACCC 368  
Db 797 TGCTACAGATGTAGTGGACCAATTTATATGTTTATGATGGGACCTCAATTTATGACCC 856  
QY 369 GCTAGTTCTGCAATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCC 428  
Db 857 TCTGATTCGCTTCTAGTGGCTCATTTGTTCTGAAAGAGATGGCAATGAGACGGCTCC 916  
QY 429 TGAGTTGTTGCCACATCAGGTTATGCTGCTGCATTTTATGATGATGCTGCTTATAA 488  
Db 917 TGAGTCTACTGCTCACTTCAAGTTATGCTGCTGCATTTTTCAGTGTGATGCTGCTTATA 976  
QY 489 TTTGACTGGATTAATATTTACTACAGTTTGTATGATGTTCCAAATTAATGCTCAGGCGG 548  
Db 977 TCTGACTGGATTAATATTTACTACATTTTACATTTTGTGATGTTCCGAAATTAATGCTCAGGCGG 1036  
QY 549 AGGAGAGTGAAGATCAGTAATAGCAGGATCTGTTGAATGTGAATGTTCTGAAAACCTG 608  
Db 1037 AGGAGAGTGAAGACAGTAACAGCAGCAGCGCTGTTGAGTGTGATGTTCTGAAAACCTG 1096  
QY 609 GAAAGGTGAAGCATGTGACATTTCTCACTGTACAGACAACTGTGTTTCTCTCATCGAGG 668  
Db 1097 GAAAGGGAGTGTGTGACATTTCTCACTGTACAGACAACTGTGTTTCTCTCATCGAGG 1156  
QY 669 CATCTGCAATTCAGATGATGTGAGAGATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGG 728  
Db 1157 CATCTGTAATCAAGACGATACAGAGGGTGTCTCTGCTTCTCTCACTGGCAGGGTCTCTGG 1216  
QY 729 ATGTTCACTTCTGACAGCTAACAGCTATTTTGGACTCGAGAGAAATATTTCAACTT 788  
Db 1217 ATGTTCAATTCCTGTCAGCTAACAGCTATTTTGGACTCGAGAGAAATATTTCTGATTT 1276  
QY 789 AAAGCTCCCGACAGCATCTCATAAAGCTGTGGTCAATGGAAACATTTATGTTGGTGTGG 848  
Db 1277 AAAGCTCCCGACAGCTCTCATAAAGCTGTGGTCAATGGAAATATAATGTTGGTGTGG 1336  
QY 849 AGGATATATGTTCAACCACTCAGATTTATACATGGTCTAGCGTATGACCTTGTCTCTAG 908  
Db 1337 CGGATATATGTTCAACCACTCAGATTTATACATGGTCTAGCGTATGACCTTGTCTCTAG 1396  
QY 909 GGAGTGGCTTCCACTTAAACCGCTTCTGTGAACAAATGTTGTAGATATGTTCAATCTTT 968  
Db 1397 GGAATGGCTTCCACTTAAACCACTTCTGTGAACAGTGTGTTGTGAAGATATGTTCAATCTTT 1456  
QY 969 GGCATATACAGGATTAATTTACATGTTATGAGGAAATTTCAATTTCACTGGAATGT 1028

Db 1457 GGCATTACATAGGATAAATCTACATGTATGGAGGAAAATGATTCAACAGGGAAGCT 1516  
QY 1029 GACCAATGAGTGGAGAGTTTTTACATATCATATAGTATCATGGTGTGTTGGACCCCTAA 1088  
Db 1517 GACCAATGAGTGGAGAGTATTTTATATATCAATATGATCATGGTATGTTTAACTCCGAA 1576  
QY 1089 GGCARAGGAGCAGTATGAGTGTGGGCACTCTGCACACATTTGTACACTGAAGAATGG 1148  
Db 1577 AGCTAAGATCATATGCAATGAGTGTGGACACTCAGACACATTTGTACACTGGCATCTGG 1636  
QY 1149 CCGAGTGGTCAATGCTGGTCAATTTTGGTCACTGCTCCCTCTCTATGGATATATAAGCAATG 1208  
Db 1637 CCGTGTGGTCAATGCTGGTCAATCTTGGTCACTGCTCCCTCTATGGATATATAAGGTTGT 1696  
QY 1209 GCAGGAATATGATTTGGATAAGACACATGGAGTATTTACACACCCAGGGTGCCTTGT 1268  
Db 1697 GCAGGAATATGATTTGGATAAGACACATGGAGTATTTACATACCTCAGGGTGCCTTGT 1756  
QY 1269 GCAAGGGGGTTACGGGCATAGCAGTGTCTTACGACCATAGGACCCAGGGCCCTATAGTTCA 1328  
Db 1757 GCAAGGGGGTTATGGCCACAGATAGTGTCTTATGATGACAGGACCAAGGCTCTGTAGTTCA 1816  
QY 1329 TGGTGGCTACAGGCTTTTCTGCTCCATTAAGTACCGCTTGCAGATGATCTTACCGATA 1388  
Db 1817 TGGTGGCTACAGGCTTTTCTGCTCCATTAAGTACCGCTTGCAGATGATCTTACCGATA 1876  
QY 1389 TGATGTGGATACCCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTGA 1448  
Db 1877 CGATGTGGATACCATGATGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTGA 1936  
QY 1449 CACAGCTGTGATAGTGGAGTGAACCATGCTGGTGTGGGGGAAACACACAAATGACAC 1508  
Db 1937 TACAGCTGTGATAGTGGAGTGAACCATGCTGGTGTGGGGGAAACACACAAATGACAC 1996  
QY 1509 ATCTATGAGCATGGCGCCAAATGCTTCTTCCAGATTTTCTGCGCTATGACATGGCTG 1568  
Db 1997 TTTCCATGAGCCAGTGGCCAAATGCTTCTTCCAGATTTTCTGCGCTATGACATGGCTG 2056  
QY 1569 TGACCGGTGTGATGCTTCCAGACCTGA - TTCCACCATGATGTCAACAGATTTGGCCA 1627  
Db 2057 TGACCGGTGTGATGCTTCCAGACCTGAAGTGTCAACAGATTTGGCCA 2116  
QY 1628 TTCCAGGATGCTTACACACAGCACCATGATGTGTTCCGTTCAATAGTCTCTCC 1687  
Db 2117 TT - CACGAGTCTTTGTACACAGCACCATGATGTGTTCCGCGCTTCAACAGCCTCTCC 2175  
QY 1688 TCAGGACATCTCTGTTATTCACCTCGGACAGTGTGATGCGCATCGGAGTGAAGCGCTT 1747  
Db 2176 TCAGTGCAGTCTTGGTCTTTACCTCGGACAGTGCATGACACACCGCATGAGTGTCTT 2235  
QY 1748 GTTTAGCAGCAGACCTGGTATTCGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCT 1807  
Db 2236 GTGTGGCAGCAGCAGCTGGTATCCGGTGTCTGTGGGACACACAGTGTCTCTCGATGTACCT 2295  
QY 1808 CGTGGCGCTGGCACTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCAAA 1867  
Db 2296 CCTGGAGTTGGCACTGAAGAAACAGAGAAAGTTAAATCAGATGTTTTTCTAAA 2355  
QY 1868 GAACCTTTGACCATGACAGATGTGACAGCAGACAGATTTGTACAGCTGCACAGCAACA 1927  
Db 2356 GAACCTTTGACCATGACAGATGTGACAGCAGACAGATTTGTACAGCTGCACAGCAACA 2415  
QY 1928 CCAA 1931  
Db 2416 CCAA 2419

## RESULT 6

US-09-893-238-12

; Sequence 12, Application US/09893238

; Patent No. US20020150973A1

; GENERAL INFORMATION:

; APPLICANT: Moore, K.

; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: PastSeq for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 6370  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-893-238-12

Query Match 37.9%; Score 1364.4; DB 9; Length 6370;  
Best Local Similarity 95.0%; Pred. No. 0;  
Matches 1492; Conservative 7; Mismatches 48; Indels 23; Gaps 8;  
QY 2035 ATGTAATCTACTTAACAAGAACACCAAGCTGCAGAGCTGTGCCCTGGACCAAGCTGCCAG 2094  
Db 1 ATGTAATCTACTTAACAAGAACACCAAGCTGCAGAGCTGTGCCCTGGACCAAGCTGCCAG 60  
QY 2095 TGGGAGCCCCGGAATCAGGAGTGCATTCCTCCCGAAATATCTGTGGCATTTGGCTGG 2154  
Db 61 TGGGAGCCCCGGAATCAGGAGTGCATTCCTCCCGAAATATCTGTGGCATTTGGCTGG 120  
QY 2155 CATTTGGTTGAAACTCATGTTTGAATAATTACTACTGCCAAGAGAAATTATGACAATGCT 2214  
Db 121 CATTTGGTTGAAACTCATGTTTGAATAATTACTACTGCCAAGAGAAATTATGACAATGCT 180  
QY 2215 AAATGTTCTGTAGAACCAACATGCCCTTTTGGCTTCTTTACAAACCAGAAAGGTA 2274  
Db 181 AAATGTTCTGTAGAACCAACATGCCCTTTTGGCTTCTTTTACAAACCAGAAAGGTA 240  
QY 2275 GAATTTGCTTAAAGCAGCTGCGAATTAATGAGTCACTCTCAGAGCATGTCCTCAAGCTCACC 2334  
Db 241 GAATTTGCTTAAAGCAGCTGCGAATTAATGAGTCACTCTCAGAGCATGTCCTCAAGCTCACC 300  
QY 2335 TTAACCCCATGGTCCGCGCTTTCGGAAGGTCAATGTGTCTTCTTCTTCTTCTGCTGGGGAAG 360  
Db 301 TTAACCCCATGGTCCGCGCTTTCGGAAGGTCAATGTGTCTTCTTCTTCTTCTGCTGGGGAAG 360  
QY 2387 AAGATATGTCCTCCATTTTACAAATAG-TTTTACTACAGTGGATGCCGCTTCTTGGAGCCAG 420  
Db 361 GATATGTCCTCCATTTTACAAATAG-TTTTACTACAGTGGATGCCGCTTCTTGGAGCCAG 420  
QY 2442 TGATGCT--GGATTTCTGTGGAATTTT-----ATCAGAACCCAGTACTCTGGGGACTGAAGG 2494  
Db 421 TGTGCTTGGATTTCTGTGGAAATTTTATTCAGGAACCCAGTACTTCTGGGGACTGAAGG 480  
QY 2495 CTGCAACCTGCAAT--CAACCCACTCAATGTTAGTGTCTGTGAAGGCTCTCAAAACACAG 2552  
Db 481 CTGCAACCTGCAAT--CAACCCACTCAATGTTAGTGTCTGTGAAGGCTCTCAAAACACAG 540  
QY 2553 TGCTAA--GCAGTGGCGGACACCATGTGCTTTGAGGACAGCATGTGGAGATTGCACCAAGCG 2611  
Db 541 TGCTAAGCAGTGGCGGACACCATGTGCTTTGAGGACAGCATGTGGAGATTGCACCAAGCG 600  
QY 2612 GCAGCTGTGATGATGTGGAGCAACATGAAGCAGTGTGTGAGCTTCCATGCTATG 2671  
Db 601 GCAGCTGTGATGATGTGGAGCAACATGAAGCAGTGTGTGAGCTTCCATGCTATG 660  
QY 2672 TGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTATACCATGAGCACTTGCCTCCCTG 2731  
Db 661 TGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTATACCATGAGCACTTGCCTCCCTG 720  
QY 2732 AAAATTGTTACGGCTACTGTACTCTGTAGTCAATGCTTTGGAGCAACACAGGCTGTGGCTGT 2791



QY 1001 GAGGAAAAATGATTCACAA--CTGGGAATGTGACCAATGAGTTGAGAGTTTTTTCACATTC 1057  
DB 1686 GAGCGAAGATTGAAACAAATGATGGCAATGTACAGATGAAATATGGGTTTTTAAACATAC 1745  
QY 1058 ATAATGAGTCATGGGTGTTGTTGACCCCTAAGCCAAAG-----AGCAGTATGCAG 1108  
DB 1746 ATAGTCAGTCATGGAGTACAAAACTCCTACTGTCTTGACATGGTCAGCAGTATGCTG 1805  
QY 1109 TGGTTGGGCACCTGACACATTTGTACACTGAAGNATGCCGAGTGGTCATGCTGGTCA 1168  
DB 1806 TGGAGGGCAATTCAGCACATATTATGGAGTTGGATAGTAGAGATGTGTGCATGATCAATA 1865  
QY 1169 TCTTTTGGTCAGTCCCTCTCTATGGATATATAAGCAATGTGAGGAATATGATTTGGATA 1228  
DB 1866 TATTTGATATTTCTGCAATATATGGTTATACAAGCAGCATACAGGAATACCATATCTCAT 1925  
QY 1229 AGAACACATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGTACGGGCATA 1288  
DB 1926 CAAACACTTGGCTTGTTCAGAAACTAAAGGAGCTATTGTACAAGGTGGATATGGCCATA 1985  
QY 1289 GCAGTGTTCAGACCATAGGACCAGGGCCCTATACGTTTCATGTTGGCTACAAGGCTTTCA 1348  
DB 1986 CTAGTGTGTATGATGAATAACAAAGTCCATTTATGTTTATGGAGGGTATAAGCATTGC 2045  
QY 1349 GTGCCAATAGTACCGCTTGCAGATGATCTCTACCGATATGATGTGGATACCGAGATG 1408  
DB 2046 CAGGGAACAAATATGGATTGGTTGATGATCTTTATAAATATGAAGTTAAACATAAGACATT 2105  
QY 1409 GGAACCTTTTAAAGACACCGATTTTCCGTACTTGCACACAGCTGTGATAGTAGTG 1468  
DB 2106 GGACTATTTTGAAGAAAGTGGGTTTGCAGATACCTTCATTACAGCTGTCTTATCAATG 2165  
QY 1469 GAAACCTGTGTTGTTGGGGAAACACACAAATGACACATCTATGAGCCATGGCGCA 1528  
DB 2166 GAGCTATGCTTATTTTGGAGGAAATACCCATATGACACTTCTTGAGTAACGGTGCAA 2225  
QY 1529 AATGCTTCTTTCAGATTTATCGGCTATGACATTTGCCTGTGACCGCTGGTCAGTGCTTC 1588  
DB 2226 AATGTTTCTGCGGATTTCCCTGGCATATGACATGCTTGTGATGAATGGNAATACTAC 2285  
QY 1589 CCAGACCTGAT-TCCACCATGATGTCAACAGATTTGGCCATTCACGAGCTCTTACACAAC 1647  
DB 2286 CAAAAACCAATCTTCATAGAGATGTCAACAGATTGGACACT-CTGCACTAGTCAATTAC 2344  
QY 1648 AGCACCATGATGTGTCGGTGTTCATATAGTCTCTCTCTCAGACACATCTCTGGTATTC 1707  
DB 2345 GGGTCCATGATATATTTGGGGGATTTCTAGTGTACTCTCTTAATGATCTCTTGTATAC 2404  
QY 1708 ACCTCGGACAGTGTGATGGCATCGAGTGAAGCCGCTTGTTTAGCAGCAGACCTGTT 1767  
DB 2405 AAGCCTCCAAATTCGAAGGCTTTCAGAGATGAAGAACTTTGTAAAAATGCTGGTCAGGG 2464  
QY 1768 ATTCGGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCGTGGCGCTGGCAACTGAT 1827  
DB 2465 ATAAATGTGTTTGGAAATAAAATCACTGTGAAT-----CTTGGGAATCTGGG 2512  
QY 1828 GAACAGACAAAGTTAAATCAGATGTGTTTCCAAAGAACTCTTGACCATGACAGA 1887  
DB 2513 AATACTAATAATATCTTAGAGCAAGTGCCTCTCTTAAACAGCTCTTCTGATGACAGA 2572  
QY 1888 TGTGACCAGCACACAGATTTGTACAGCTGCACAGCCAAACACCAATGACTGCCACTGGTGC 1947  
DB 2573 TGTTACAGATATGCAGATGTGCCAGCTGTACTTGCCCAATACAAATGGGTGCCAATGGTGT 2632  
QY 1948 AATGACCAATTTGTGCCCAAGAACACAGCTGTCTCAGAGGCCAGATCTCCATTTTATAG 2007  
DB 2633 GATGACAAGAAATGCAATTTCCGGCAATAGTAAC-----GCAGTATGTCTGTCAAGAAC 2686  
QY 2008 TATGAGAAATGCCCCAGGATAACCTTATGTACTGTACTGTAAACAAGAAAGACCGCTCAGG 2067  
DB 2687 TACACCAATATGATGTAGAAA---TGACAGATTTGTAAACAACTTACCAGCTGTAAA 2743  
QY 2068 AGCTGTGCCCTGGACCAAGAACTGCCAGTGGGAGCCCCGGGAATCAGGAGTGCATTTGCCCTG 2127

DB 2744 AGCTGTTCACTAAACTTGAATTCAGATGGGATCAGAGACAGCAAGAATGCCAGGCTTTA 2803  
QY 2128 CCGGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACACTCATGTTTGAATAATCT 2187  
DB 2804 CCAGCTCATCTTTGTGGAGAGGATGGAGTCATATTGGGATGCTTGTCTTAGAGTCAAT 2863  
QY 2188 ACTCCAAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACCAACAATGCCCTTTTG 2247  
DB 2864 TCCAGTAGAGAAACTATGACAAATGCAAACTTTATTCCTATATCTTAGTGGAAATCTT 2923  
QY 2248 GCTTCTCTTCAACCCCAAGAAAGGTAGAAATTTCTCTTAAGCAGCTGCGGAATAATGCAG 2307  
DB 2924 GCTTCATTAACAACCTCAAAAGAAAGTAGAATTTGTTTC-----TGCAAT 2965  
QY 2308 TCATCTCAGAGCATGTCACAGCTCACCTTAACCCCATGGGTGCGCCTTCGGAAGATCAAT 2367  
DB 2966 GAAATACAGAAAGTATACACACAGAAAGTATCACCTTGGTAGGCTTGGCGAAGTCAAT 3025  
QY 2368 GTGTCTCTA CTGGTCTCTGGGAGATATGTCCCAATTTTACAATAGTTTACTACAGTGGATG 2427  
DB 3026 ATATCTTATTGGGATGGGAGACATGCTCTCTTTACAAACACACACTACAGTGGCTT 3085  
QY 2428 CCGTCTGAGCCCACTGATGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGA 2487  
DB 3086 CCTGGCGAAACCAATGATTTCTGGGTTTTGTGCAATCTGGAAGGGCTGCACTGGCAGGC 3145  
QY 2488 CTGAAGGCTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTCTG----- 2542  
DB 3146 TTAAAGGCTAATCTTGTACATCTATGGCAATGGCCTTGTCTGTGAAGAAACCTGTGTT 3205  
QY 2543 ----CAAACCACTGCTAGCAGTGGCGGACACCATGTGCTTGGAGCAGCAGTATGGGA 2598  
DB 3206 AGTCCAAATCAAAATGGAGCCGTGCAAAAGCCATGCTCTCTGAGGACATCATGTCTCC 3265  
QY 2599 GATTGACACCGCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAGAGCAGTGTGGAC 2658  
DB 3266 AACTGTACAACAAATGGCATGAGTGTATGTGTGTCAGCAGTACGAACGATGTGTGAC 3325  
QY 2659 TCCAATGCTATGGGCTCTCTCTCTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2718  
DB 3326 TCTAATGCTATATCATCTCTTTCCATATGGAACAATGTCTAGAGTGGCAAC-----TGCC 3382  
QY 2719 ACCTGCCCCCTGAAATTTGTTTCAGGCTACTGTACCTGTAGTCTGCTGGAGCAACCA 2778  
DB 3383 ACCTGCTCCCTCTAAATTTGTTCTGGATGTGAACCTGTGGACAGTGTGTTGGAACAGCT 3442  
QY 2779 GGCTGTGCTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCTATAGAGGGTTCC 2838  
DB 3443 GGATGTGCTGTGCAATGATCCTAGTAATAACAGGAAGAGGACATTTGCATTGAAGGTTCT 3502  
QY 2839 TATAAGACACAGTGAAGATGCTTGGCAAGCCCTTACAGGAATTTCTATCCACAGCCC 2898  
DB 3503 TCACGGGACCAATGAAG-----CTTATTTGGAATGCCAACCAATGAGATG 3547  
QY 2899 GTGCTCAATTCAGCATGTGTCTAGAGCAGCAGATACAACTGCTCTTTTCATTCACCTGT 2958  
DB 3548 GTTCTTGACCAATCTTTGGCCCCCAAGAAAGAACTATGAGTGTCTTTATCCAGTGT 3607  
QY 2959 CCAGCTTCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGT 3018  
DB 3608 CCAGCTTCCAGTGTAAATGGACATAGCAGCTTGCATCAATAATAATGTGTGCGAAGTGT 3667  
QY 3019 GAGAACCTTGACCAAGCAAGCAGCTTGGAGACCTGTCATATCTGCTTCTACGGTGAATCCC 3078  
DB 3668 AAAAATCTCACACAGGAAGCAGTGTCAAGTTGTATGCCAGGTTATTATGGAGATCCA 3727  
QY 3079 ACCAATGGAGGAAATGTGAGCCATGCAAGTGAATGGGACGCGTCTCTGTGCAACACC 3138  
DB 3728 ACCAATGTGGACAGTGCACAGCTTGTACATGCACTGGCCATGCAAAATCTGTCTCATCTG 3787  
QY 3139 AACAGGGCAAGTGTCTTCTGCAACCAAGGGCGTCAAGGGGAGGAGTGCACGCTATGT 3198



Db 3788 CACACAGGAAATGTTTCTCGCACAACTAAAGGAATAAAGGTGACCAATGCCAATATATGT 3847  
Qy 3199 GAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATATCTTCTT 3258  
Db 3848 GACTCTGAAAATCGTATGTTGGTAAATCCACTTAGAGGAACATGTTATTACAGCCTTTGG 3907  
Qy 3259 ATTGACTATCAGTTCACCTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGTATC 3318  
Db 3908 ATTGAATTATCAATTTACCTTCAGCTTATTACAGAGAGATGATCGCCACCACTATGCCATA 3967  
Qy 3319 AATTTTCTGGCTACTCTCGTACGACAAACAGGATTTGACATGTTTCATCAATGCTCC 3378  
Db 3968 AACTTTATAGCAACCCAGACAGTCCGACAAAATCTGGATATATCAATTAATGATCA 4027  
Qy 3379 AAGATTTCAACCTCAACATCAGCTGGGTGCGCAGTTTCTCAGTCCGAAACCCAGGCTGA 3438  
Db 4028 AACAACTTTAATCTCAACATTTCTGCTGTCTGCTGTTCAACAGCTGGAACAATATCTGGG 4087  
Qy 3439 GAAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAT 3498  
Db 4088 GAAGAGACTTCTATGTTTCCAGATATATTAAGAAATACAGATAGTTTCTCTAT 4147  
Qy 3499 GAGAGTTTGAATTTCCGACACCCCAATATCACTTTCTGTTTATGTCAAGTAAATTC 3558  
Db 4148 GAAAATTTAACTTTAGAGCAATCTCAACATTTACATTTCTATGTACGTCAAGCACTTT 4207  
Qy 3559 ACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3594  
Db 4208 TCCTGGCCTATTAAATACAGATGCAATTCACAA 4243

## RESULT 8

US-09-893-238-10  
; Sequence 10, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 1051  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-10

Query Match 12.9%; Score 463; DB 9; Length 1051;  
Best Local Similarity 85.2%; Pred. No. 1.3e-132;  
Matches 517; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 15 GCGGGCACTGAGCAAGGTCTGAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 74  
Db 441 GAGGGCGGAGCGGCGGCTGCGGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 500  
Qy 75 GAGCGGCGGCGGCGGCGGCTGCTCAACCGCGGTCTGCAACCCCTGGCACCGGCGGCGGCGGT 134  
Db 501 GGAATGTGACCGGCGGCTGTCTCAACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGT 560  
Qy 135 CTGCCCCCGCGGCTGGGTGGCGGAGCAATGCCAGCACTGCGGGGCGGCGGCGGCGGCGGCGG 194  
Db 561 CTGCCCCCGCGGCTGGGTGGCGGAGCAATGCCAGCACTGCGGGGCGGCGGCGGCGGCGGCGG 620

Qy 195 TGGATCTTCTGGGTTTGTGACAGATGACCTGGAATTTATAATACAAAACGAGTGCAC 254  
Db 621 TGGCTCTTCTGGATTTGTAAACAGATGGAACCTGGGAATTTATAAATATAAGACGAAGTGCAC 680  
Qy 255 GTGGCTCATTCAGAGCAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTAC 314  
Db 681 ATGGCTCATTCAGAGCAGCCAAATAGAAATAGAGACTTCGTTTCAACCATTTTGTCTAC 740  
Qy 315 AGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGT 374  
Db 741 AGAATGTAGCTGGGACCATTTATATGTTTATGATGGGACTCAATCTACGACCTCTGAT 800  
Qy 375 TGCTGCATTTAGTGCCTTCATTTGCTGAGAGATGGCAATGAGACTGTCCTCTGAGGT 434  
Db 801 TGCTGCCTTTAGTGCCTTCATTTGCTGAAAGAGATGGCAATGAGAGCGCTCTCTGAGGT 860  
Qy 435 TGTTCGCCACATCAGGTTATGCTTTCCTGTCATTTTTTTTATGATGCTCTTATATTTGAC 494  
Db 861 CACTGTCACTTCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 920  
Qy 495 TGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 554  
Db 921 TGGATTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 980  
Qy 555 GTGTAAGATCAATAGCAGCGATCTGTTGATGTCAATGTTTCTGAAAACCTGGAAGG 614  
Db 981 GTGTAAGAGCAGTAACAGCAGCGGCTGTTGAGTGTGAATGTTTCTGAAAACCTGGAAGG 1040  
Qy 615 TGAAGCA 621  
Db 1041 GCGCGGA 1047

## RESULT 9

US-10-198-846-5869/c  
; Sequence 5869, Application US/10198846  
; Publication No. US20030099974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steinmann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5869  
; LENGTH: 625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-5869

Query Match 7.7%; Score 277.4; DB 14; Length 625;  
Best Local Similarity 99.6%; Pred. No. 6.4e-75;  
Matches 278; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 2477 GTACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCACTCAATGTGTGTGTGAAA 2536  
Db 340 GTACTCGGGGACTGAAGGCTGCAACCTGCAATCAACCACTCAATGTGTGTGTGAAA 281  
Qy 2537 GGCTTGCAACCAAGTGTGAAGCAGTGGCGGACCACTGCTGCTGAGCAGCATGTG 2596  
Db 280 GGCCTGCAACCAAGTGTGAAGCAGTGGCGGACCACTGCTGCTGAGCAGCATGTG 221



```
QY 2597 GAGATTGACCCAGCGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGG 2656
Db 220 GAGATTGACCCAGCGCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGG 161
QY 2657 ACTCCAAATGCTATGCGCTCTCCCTTTGGCCAGGTATGAATGTTATAGATGA 2716
Db 160 ACTCCAAATGCTATGCGCTCTCCCTTTGGCCAGGTATGAATGTTATAGATGA 101
QY 2717 GCACCTGCCCCCTGAAATTTGTTACAGGCTACTGTACCT 2755
Db 100 GCACCTGCCCCCTGAAATTTGTTACAGGCTACTGTACCT 62

RESULT 10
US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonmax Sequence Listing Engine vers. 1.1
; SEQ ID NO 15900
; LENGTH: 531
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.14
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3
```

```
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9
US-09-864-761-15900
Query Match 7.0%; Score 250.8; DB 9; Length 531;
Best Local Similarity 97.3%; Pred.No.1.1e-66;
Matches 255; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY 973 TTATACAGGATAAAATTTACATGATGAGGAGGAAAAATTTGATTCACACTGGGAATGTGACC 1032
Db 141 TTTTCCTAGGATAAAATTTACATGATGAGGAGGAAAAATTTGATTCACACTGGGAATGTGACC 200
QY 1033 AATGAGTTGAGAGTTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCA 1092
Db 201 AATGAGTTGAGAGTTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCA 260
QY 1093 AAGGAGCAGTATGAGTGGTGGGCACTCTGCAACAATTTGTTACACTGAAGAATGGCGGA 1152
Db 261 AAGGAGCAGTATGAGTGGTGGGCACTCTGCAACAATTTGTTACACTGAAGAATGGCGGA 320
QY 1153 GTGGTCATGCTGGTCATCTTTTGGTCACTGGCCCTCTCTATGGATATATAAGCAATGTGCAG 1212
Db 321 GTGGTCATGCTGGTCATCTTTTGGTCACTGGCCCTCTCTATGGATATATAAGCAATGTGCAG 380
QY 1213 GAATATGATTTGGATAAGACA 1234
Db 381 GAATATGATTTGGGTAGGTATA 402

RESULT 11
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. US20020168637A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darriack
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; APPLICANT: Wang, Aijun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966
Query Match 6.8%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred.No.1.9e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2293 CTGGGAATAATGAGTCATCTCAGAGCATGTCACCTTAACCCCATGGGTGGC 2352
Db 246 CTGGGAATAATGAGTCATCTCAGAGCATGTCACCTTAACCCCATGGGTGGC 187
QY 2353 CTTCGGAAGATCAATGTGCTCTACTGTGGTGGGAAGATGTCCCCATTTACAAATAGT 2412
Db 186 CTTCGGAAGATCAATGTGCTCTACTGTGGTGGGAAGATGTCCCCATTTACAAATAGT 127
QY 2413 TTACTACAGTGGATGCCCTCTGAGCCCAAGTCATGCTGGATTTTATCAGAA 2472
Db 126 TTACTACAGTGGATGCCCTCTGAGCCCAAGTCATGCTGGATTTTATCAGAA 67
QY 2473 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGTCATCAACCCACTCAATGTTGTCTGT 2532
```

```
Db 66 CCCAGTACTCGGGAGCTGAAGGCTGCACCTGCAACCTGCATCAACCACTCAATGCTAGTGTCTGT 7
QY 2533 GAAAGG 2538
Db 6 GAAAGG 1

RESULT 12
US-09-902-941-966/c
; Sequence 966, Application US/09902941
; Patent No. US20020172952A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-966

Query Match 6.8%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.9e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 CTGCGAATAATGCAGTCAATCTCAGAGCATGTCCAAAGCTCACTTAACCCCATGGGTCGGC 2352
Db 246 CTGCGAATAATGCAGTCAATCTCAGAGCATGTCCAAAGCTCACTTAACCCCATGGGTCGGC 187
QY 2353 CTTCCGAGATCAATGTCTCTACTGCTGCGGAGAGATATGTCGCCATTTTATCAGAA 2472
Db 186 CTTCCGAGATCAATGTCTCTACTGCTGCGGAGAGATATGTCGCCATTTTATCAGAA 127
QY 2413 TTACTACAGTGGATGCCCTCTGAGCCCAAGTCACTGCTGGATTCCTGGAAATTTTATCAGAA 2472
Db 126 TTACTACAGTGGATGCCCTCTGAGCCCAAGTCACTGCTGGATTCCTGGAAATTTTATCAGAA 67
QY 2473 CCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCACTCAATGCTAGTGTCTGT 2532
Db 66 CCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCACTCAATGCTAGTGTCTGT 7

QY 2533 GAAAGG 2538
Db 6 GAAAGG 1

RESULT 14
US-10-017-754-966/c
; Sequence 966, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnerakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-966

Query Match 6.8%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.9e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 CTGCGAATAATGCAGTCAATCTCAGAGCATGTCCAAAGCTCACTTAACCCCATGGGTCGGC 2352
```

```
Db      246  CTGCGAATATGAGTCATCTCAGAGCATGTCAAGCTCACCTTAACCCATGGGTGGC 187
QY      2353  CTTTGGAAAGATCAATGTGTCCTACTGGTCTGGGAAGATATGTCCCAATTTACAAATAGT 2412
Db      186  CTTTGGAAAGATCAATGTGTCCTACTGGTCTGGGAAGATATGTCCCAATTTACAAATAGT 127
QY      2413  TTACTACAGTGGATGCGGTCTGAGCCCACTGATGCTGGATTCTGTGGAAATTTTATCAGAA 2472
Db      126  TTACTACAGTGGATGCGGTCTGAGCCCACTGATGCTGGATTCTGTGGAAATTTTATCAGAA 67
QY      2473  CCCAGTACTCGGGACTGAAGGTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 2532
Db      66  CCCAGTACTCGGGACTGAAGGTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 7
QY      2533  GAAAGG 2538
Db      6  GAAAGG 1
```

## RESULT 15

```
US-10-113-872-966/c
; Sequence 966, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-966
```

```
Query Match      6.8%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 1.9e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2293  CTGCGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCATGGGTGGC 2352
Db      246  CTGCGAATATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCATGGGTGGC 187
QY      2353  CTTTGGAAAGATCAATGTGTCCTACTGGTCTGGGAAGATATGTCCCAATTTACAAATAGT 2412
Db      186  CTTTGGAAAGATCAATGTGTCCTACTGGTCTGGGAAGATATGTCCCAATTTACAAATAGT 127
QY      2413  TTACTACAGTGGATGCGGTCTGAGCCCACTGATGCTGGATTCTGTGGAAATTTTATCAGAA 2472
Db      126  TTACTACAGTGGATGCGGTCTGAGCCCACTGATGCTGGATTCTGTGGAAATTTTATCAGAA 67
QY      2473  CCCAGTACTCGGGACTGAAGGTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 2532
Db      66  CCCAGTACTCGGGACTGAAGGTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGT 7
QY      2533  GAAAGG 2538
Db      6  GAAAGG 1
```

Search completed: March 5, 2004, 13:18:25  
Job time : 802.175 secs

**This Page Blank (uspto)**



Db 534 AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCAACCGTAGT 593  
Qy 375 TGCTGCATTTAGTGGCCCTCAATGTTCTGTGAGAGATGGCAATGAGACTGTCCCTCAGGT 434  
Db 594 TGCTGCATTTAGTGGCCCTCAATGTTCTGTGAGAGATGGCAATGAGACTGTCCCTCAGGT 653  
Qy 435 TGTTGCCACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTGAC 494  
Db 654 TGTTGCCACATCAGGTTATGCTTGTGCTGCAATTTTATGATGCTGCTTATATTTGAC 713  
Qy 495 TGGATTTAATATTTACTTTACAGTTTGTATGTTGTGTCMAATPAATGCTCAGGCCGAGAGA 554  
Db 714 TGGATTTAATATTTACTTTACAGTTTGTATGTTGTGTCMAATPAATGCTCAGGCCGAGAGA 773  
Qy 555 GTGTAAAGATCAGTAATAGCAGCGATCTGTTGTAATGTGAATGTTCTGMAAATCGAAAGG 614  
Db 774 GTGTAAAGATCAGTAATAGCAGCGATCTGTTGTAATGTGAATGTTCTGMAAATCGAAAGG 833  
Qy 615 TGAAGCATGTGACATTCCTCACTGTACAGACAACCTGTGTTTCTCATCGAGGCATCTG 674  
Db 834 TGAAGCATGTGACATTCCTCACTGTACAGACAACCTGTGTTTCTCATCGAGGCATCTG 893  
Qy 675 CAATTCAGTGTGATGTGAGAGATGCTCCTGCTTCTCAGACTGCGAGGTCCTGATGTTG 734  
Db 894 CAATTCAGTGTGATGTGAGAGATGCTCCTGCTTCTCAGACTGCGAGGTCCTGATGTTG 953  
Qy 735 AGTTCTCTGTACCACTAACAGTCAATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCT 794  
Db 954 AGTTCTCTGTACCACTAACAGTCAATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCT 1013  
Qy 795 CCCAGAGCATCTCATAAAGCTGTGTTCAATGGAACATTTATGTGGTGTGAGAGATA 854  
Db 1014 CCCAGAGCATCTCATAAAGCTGTGTTCAATGGAACATTTATGTGGTGTGAGAGATA 1073  
Qy 855 TATGTTCAACCACTCAGATTAATCACTGTTCTAGCGTATGACCTTGTCTTAGGGAGTG 914  
Db 1074 TATGTTCAACCACTCAGATTAATCACTGTTCTAGCGTATGACCTTGTCTTAGGGAGTG 1133  
Qy 915 GCTTCCACTAAACCGTCTGTGAAACAAATGTGGTGTGTAGATATGTCATTTTGGCATT 974  
Db 1134 GCTTCCACTAAACCGTCTGTGAAACAAATGTGGTGTGTAGATATGTCATTTTGGCATT 1193  
Qy 975 ATACAAGGATAAATTTACATGATGAGGAGAAATTTGATTCAACTGGGAATGTGACCAA 1034  
Db 1194 ATACAAGGATAAATTTACATGATGAGGAGAAATTTGATTCAACTGGGAATGTGACCAA 1253  
Qy 1035 TGAGTTGAGAGTTTTCACATTCATAATGAGTCAATGGTGTGTTGACCCCTAAGGCAAA 1094  
Db 1254 TGAGTTGAGAGTTTTCACATTCATAATGAGTCAATGGTGTGTTGACCCCTAAGGCAAA 1313  
Qy 1095 GGAGCATGTGAGTGTGGGCACTCTGCACACATTTGTTACACTGGAATGCGCGAGT 1154  
Db 1314 GGAGCATGTGAGTGTGGGCACTCTGCACACATTTGTTACACTGGAATGCGCGAGT 1373  
Qy 1155 GGTCAATGCTGTCTATTTGGTCACTGCCCTCTCTATGAGATATATAAGCAATGTGACGA 1214  
Db 1374 GGTCAATGCTGTCTATTTGGTCACTGCCCTCTCTATGAGATATATAAGCAATGTGACGA 1433  
Qy 1215 ATATGATTTGGATTAAGAACACATGGAGTATATTACACACCCAGGTCCTGTGCAAGG 1274  
Db 1434 ATATGATTTGGATTAAGAACACATGGAGTATATTACACACCCAGGTCCTGTGCAAGG 1493  
Qy 1275 GGGTTAGCGGCATAGCAGTGTGTTACGACCATAGGACCCAGGCGCTTATACGTTTATGGTGG 1334  
Db 1494 GGGTTAGCGGCATAGCAGTGTGTTACGACCATAGGACCCAGGCGCTTATACGTTTATGGTGG 1553  
Qy 1335 CTACAAGGCTTTAGTCCCAATAGTACCGGCTGTGAGATGATCTCTACCGATATGATGT 1394  
Db 1554 CTACAAGGCTTTAGTCCCAATAGTACCGGCTGTGAGATGATCTCTACCGATATGATGT 1613  
Qy 1395 GGATACCCAGATGTGGAACCAATCTTAAGGACAGCGGATTTTTCGGTTACTTGCACACAGC 1454  
Db 1614 GGATACCCAGATGTGGAACCAATCTTAAGGACAGCGGATTTTTCGGTTACTTGCACACAGC 1673

Qy 1455 TGTGATAGTGTGAGTGAACCATGCTGGTGTGTTTGGGGGAAACACACAAATGACACATCTAT 1514  
Db 1674 TGTGATAGTGTGAGTGAACCATGCTGGTGTGTTTGGGGGAAACACACAAATGACACATCTAT 1733  
Qy 1515 GAGCCATGGCCCAAAATGCTTCTCTTCAAGATTTTCAAGCCCTATGACATTTGCTGTGACCG 1574  
Db 1734 GAGCCATGGCCCAAAATGCTTCTCTTCAAGATTTTCAAGCCCTATGACATTTGCTGTGACCG 1793  
Qy 1575 CTGGTCACTGCTTCCAGACCTGAT-TCCACATGATGTCAACAGATTTGGCCATTCAG 1633  
Db 1794 CTGGTCACTGCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAG 1682  
Qy 1634 CAGTCTTACACAAACAGCACCATGATGTTTGGTGGTTCATATGATCTCTCTCTCAGCG 1693  
Db 1853 CAGTCTTACACAAACAGCACCATGATGTTTGGTGGTTCATATGATCTCTCTCTCAGCG 1912  
Qy 1694 ACATCCTGTTATTCACCTCGGAAACAGTGTGATGGCATCGGAGTGAAGCCGTTGTTAG 1753  
Db 1913 ACATCCTGTTATTCACCTCGGAAACAGTGTGATGGCATCGGAGTGAAGCCGTTGTTAG 1972  
Qy 1754 CAGCAGGACCTGGTATTCGGTGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCTGGG 1813  
Db 1973 CAGCAGGACCTGGTATTCGGTGTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCTGGG 2032  
Qy 1814 CGTGGCACTGTGATGAACAGAGAAAGTTAAATCAGATGTTTTCGAAAGACCTC 1873  
Db 2033 CGTGGCACTGTGATGAACAGAGAAAGTTAAATCAGATGTTTTCGAAAGACCTC 2092  
Qy 1874 TTGACCATGACAGATGTGACAGACACAGATTTGTTACAGCTGCACAGCCAAACCAATG 1933  
Db 2093 TTGACCATGACAGATGTGACAGACACAGATTTGTTACAGCTGCACAGCCAAACCAATG 2152  
Qy 1934 ACTGCACTGTGTGCAATGACATTTGTTGCCAGAAACACAGCTGTCTAGAGGCCAGA 1993  
Db 2153 ACTGCACTGTGTGCAATGACATTTGTTGCCAGAAACACAGCTGTCTAGAGGCCAGA 2212  
Qy 1994 TCTCCATTTTATGATGATGAGAAATGCCCCAAGGATTAACCTTATCTACTGTAAACAAGA 2053  
Db 2213 TCTCCATTTTATGATGATGAGAAATGCCCCAAGGATTAACCTTATCTACTGTAAACAAGA 2272  
Qy 2054 AGACCAGCTGCAGGAGCTGTGCTTCCAGCAAGAACTGCCAGTGGAGCCCGGAAATCAGG 2113  
Db 2273 AGACCAGCTGCAGGAGCTGTGCTTCCAGCAAGAACTGCCAGTGGAGCCCGGAAATCAGG 2332  
Qy 2114 AGTGCATTTGCCCTGCCGAAATATCTGTGCAATTTGGCTGCTTGGTGGAACTCAT 2173  
Db 2333 AGTGCATTTGCCCTGCCGAAATATCTGTGCAATTTGGCTGCTTGGTGGAACTCAT 2392  
Qy 2174 GTTTGAAATTTACTACTGCCAAGGAGAAATTTATGACAAATGCTTAAATTTCTGTAGGAACC 2233  
Db 2393 GTTTGAAATTTACTACTGCCAAGGAGAAATTTATGACAAATGCTTAAATTTCTGTAGGAACC 2452  
Qy 2234 ACAATGCCCTTTGGCTTCTTCAACCCAGAGAGAGATGATTTGCTTCCCTTAAGCAGC 2293  
Db 2453 ACAATGCCCTTTGGCTTCTTCAACCCAGAGAGAGATGATTTGCTTCCCTTAAGCAGC 2512  
Qy 2294 TGCGAATTAATCAGTCACTCTCAGAGCATGTCACCAAGCTCACCTTAACCCCATGGTCCGCC 2353  
Db 2513 TGCGAATTAATCAGTCACTCTCAGAGCATGTCACCAAGCTCACCTTAACCCCATGGTCCGCC 2572  
Qy 2354 TFCGGAAGATCAATGTGTCCTACTGCTGGGAGATATGTCGCCATTTTACAAATAGTT 2413  
Db 2573 TFCGGAAGATCAATGTGTCCTACTGCTGGGAGATATGTCGCCATTTTACAAATAGTT 2632  
Qy 2414 TACTACATGATGCTGCTCAGCCAGTGTGATGCTGGAATCTGTGGAAATTTTATCAGAAC 2473  
Db 2633 TACTACATGATGCTGCTCAGCCAGTGTGATGCTGGAATCTGTGGAAATTTTATCAGAAC 2692  
Qy 2474 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTGTGCTGTG 2533  
Db 2693 CCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTGTGCTGTG 2752

```
QY 2534 AAGGGCTGCAAAACCAAGTCGTAAGCAGATGCGCGACACCATGTGCTTGGAGACAGCAT 2593
|
|
|
Db 2753 AAGGGCTGCAAAACCAAGTCGTAAGCAGATGCGCGACACCATGTGCTTGGAGACAGCAT 2812
|
|
|
QY 2594 GTGGAGATTGCACACAGCGGAGCTCTGAGTCATGTGGTGACACACATGAACATGTG 2653
|
|
|
Db 2813 GTGGAGATTGCACACAGCGGAGCTCTGAGTCATGTGGTGACACACATGAACATGTG 2872
|
|
|
QY 2654 TGAAGTCAATGCTATGTGGCTCTTCCCTTTCGCGAGTGTATGGAATGATACGA 2713
|
|
|
Db 2873 TGAAGTCAATGCTATGTGGCTCTTCCCTTTCGCGAGTGTATGGAATGATACGA 2932
|
|
|
QY 2714 TGAAGTCAATGCTATGTGGCTCTTCCCTTTCGCGAGTGTATGGAATGATACGA 2773
|
|
|
Db 2933 TGAAGTCAATGCTATGTGGCTCTTCCCTTTCGCGAGTGTATGGAATGATACGA 2992
|
|
|
QY 2774 AAGCAGCTGTGGCTGTGCTGATCCACAGCAATATGCGAAAGGAAATGCATAGAG 2833
|
|
|
Db 2993 AAGCAGCTGTGGCTGTGCTGATCCACAGCAATATGCGAAAGGAAATGCATAGAG 3052
|
|
|
QY 2834 GTTCCTATAAGACCGATGAGATGCTTTCGCAAGCCCTACAGGAATTTCTATCCAC 2893
|
|
|
Db 3053 GTTCCTATAAGACCGATGAGATGCTTTCGCAAGCCCTACAGGAATTTCTATCCAC 3112
|
|
|
QY 2894 AGCCCTGTCTCAATTCAGCATGTGTCTAGAGCAGCAGATACAACTGCTTTTCATTC 2953
|
|
|
Db 3113 AGCCCTGTCTCAATTCAGCATGTGTCTAGAGCAGCAGATACAACTGCTTTTCATTC 3172
|
|
|
QY 2954 ACTGTCCAGTTTCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAG 3013
|
|
|
Db 3173 ACTGTCCAGTTTCCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAG 3232
|
|
|
QY 3014 AGTGTGAGACCTGACACAGGACAGCAGTCTGAGACCTGCAATCTGCGCTTCTACGGTG 3073
|
|
|
Db 3233 AGTGTGAGACCTGACACAGGACAGCAGTCTGAGACCTGCAATCTGCGCTTCTACGGTG 3292
|
|
|
QY 3074 ATCCCAACCAATGAGGGAATGTGAGCAGTGTCAAGTGCAATGGGCGCGCTCTGTGCA 3133
|
|
|
Db 3293 ATCCCAACCAATGAGGGAATGTGAGCAGTGTCAAGTGCAATGGGCGCGCTCTGTGCA 3352
|
|
|
QY 3134 ACACCAACAGGCGAAGTGTCTGACACACAGGCGCTCAAGGGGCGAGTGCACG 3193
|
|
|
Db 3353 ACACCAACAGGCGAAGTGTCTGACACACAGGCGCTCAAGGGGCGAGTGCACG 3412
|
|
|
QY 3194 TATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATATATCTC 3253
|
|
|
Db 3413 TATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATATATCTC 3472
|
|
|
QY 3254 TTCTTATGACTATCAGTTTCACTTATCCAGGAAGATGATCGCTATATACAG 3313
|
|
|
Db 3473 TTCTTATGACTATCAGTTTCACTTATCCAGGAAGATGATCGCTATATACAG 3532
|
|
|
QY 3314 CTATCAATTTTGGCTTACTCTGACGACAAACAGGGAATTTGGACATGTTTCATCAATG 3373
|
|
|
Db 3533 CTATCAATTTTGGCTTACTCTGACGACAAACAGGGAATTTGGACATGTTTCATCAATG 3592
|
|
|
QY 3374 CCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3433
|
|
|
Db 3593 CCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGG 3652
|
|
|
QY 3434 CTGGAGAGAGATGCGCTGTGTTTCAAAACCAACATTAAGGATCAAGATAGTTTCT 3493
|
|
|
Db 3653 CTGGAGAGAGATGCGCTGTGTTTCAAAACCAACATTAAGGATCAAGATAGTTTCT 3712
|
|
|
QY 3494 CTAAATGAGAAGTTGATTTTCGCAACCCCAATATCACTTCTTCTTGTATGTGATGTA 3553
|
|
|
Db 3713 CTAAATGAGAAGTTGATTTTCGCAACCCCAATATCACTTCTTCTTGTATGTGATGTA 3772
|
|
|
QY 3554 ATTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAAACAATGA 3597
|
|
|
Db 3773 ATTTCACTGGCCCATCAAAATTCAGGTGCAAACTGAAACAATGA 3816
|
|
|
```

## RESULT 2

```
US-09-245-041-14
; Sequence 14, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 8589
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-14
```

```
Query Match 96.7%; Score 3476.8; DB 3; Length 8589;
Best Local Similarity 98.8%; Pred. No. 0;
Matches 3524; Conservative 0; Mismatches 42; Indels 2; Gaps 2;
```

```
QY 15 GCGGCCAACTCAGGCAAGGCTGAGAGAGAGGACCGCGCGACGCGAGCGCTCGCGGGCAG 74
|
|
|
Db 234 GGAGCCCGAGCGCGCTGCGGTGGCGGGCGCGGTGTCGCGTCCGCGCGAGCGAGGCCAA 293
|
|
|
QY 75 GAGCGCGGGCGCGCACTGTGTCAACGGCGGTGCTGCAACCCCTGSCACCGCGCCAGTGGT 134
|
|
|
Db 294 GGAATGTGACCGCGCGTGTGTCAACGGCGGTGCTGCAACCCCTGSCACCGCGCCAGTGGT 353
|
|
|
QY 135 CTGCCCCCGCGCTGGGTGGCGGAGCAATGCCAGCACTGCGGGGGCGCGCTTCAGACTAAC 194
|
|
|
Db 354 CTGCCCCCGCGCTGGGTGGCGGAGCAATGCCAGCACTGCGGGGGCGCGCTTCAGACTAAC 413
|
|
|
QY 195 TGGATCTTCTGGGTTTGTGACAGATGAGACCTGGGAATTAATAACAAACGAAGTGCAC 254
|
|
|
Db 414 TGGATCTTCTGGGTTTGTGACAGATGAGACCTGGGAATTAATAACAAACGAAGTGCAC 473
|
|
|
QY 255 GTGGCTCATTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTAC 314
|
|
|
Db 474 GTGGCTCATTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTAC 533
|
|
|
QY 315 AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGGAATCAATTTATGACCGCTAGT 374
|
|
|
Db 534 AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGGAATCAATTTATGACCGCTAGT 593
|
|
|
QY 375 TGCTGCATTTAGTGGCCCTCATTTGTTTCTGAGAGAGATGGCAATCAGACTGTCCTGAGGT 434
|
|
|
Db 594 TGCTGCATTTAGTGGCCCTCATTTGTTTCTGAGAGAGATGGCAATCAGACTGTCCTGAGGT 653
|
|
|
QY 435 TGTTGCCACATCAGGTATGCTTTCGTCATTTTTTTTAGTGATGCTGTTTATAATTTGAC 494
|
|
|
Db 654 TGTTGCCACATCAGGTATGCTTTCGTCATTTTTTTTAGTGATGCTGTTTATAATTTGAC 713
|
|
|
QY 495 TGGATTTAATATTACTTACAGTTTGTATGTTGTCCTCAATTAACCTGCTCAGCGCGAGGA 554
|
|
|
Db 714 TGGATTTAATATTACTTACAGTTTGTATGTTGTCCTCAATTAACCTGCTCAGCGCGAGGA 773
|
|
|
QY 555 GTGTAAGATCAGTAAATAGCAGCGATACCTTGTGAATGTGAATGTTCTGAAAACTGGAAGG 614
|
|
|
Db 774 GTGTAAGATCAGTAAATAGCAGCGATACCTTGTGAATGTGAATGTTCTGAAAACTGGAAGG 833
|
|
|
QY 615 TGAAGATGTGCAATTCCTCTCTAGACAACTGTGGTTTTTCTCTCATCAGGCACTGTG 674
|
|
|
Db 834 TGAAGATGTGCAATTCCTCTCTAGACAACTGTGGTTTTTCTCTCATCAGGCACTGTG 893
|
|
|
QY 675 CAATTTCAAGTGTGTCAGAGGATGCTCTCTGTTTCTCAGACTGGCAGGGTCTCTGATGTTT 734
|
|
|
```





2894 AGCCCTGCTCAATTCAGCATGTCTAGAGACAGCAGATCAAACTGGTCTTTTCATTC 2953  
Db AGCCCTGCTCAATTCAGCATGTCTAGAGACAGCAGATCAAACTGGTCTTTTCATTC 3172  
Qy ACTGTCAGCTTGCCTCAATGCAAGCGGCACAGTAAATGCAATCAATCAGAGCATCTGTGAGA 3013  
Db ACTGTCAGCTTGCCTCAATGCAAGCGGCACAGTAAATGCAATCAATCAGAGCATCTGTGAGA 3232  
Qy AGTGTGAGAACCTTGACACAGGGAAGCACTGCGAGACCTCATATCTGGCTTCTACGGTG 3073  
Db AGTGTGAGAACCTTGACACAGGGAAGCACTGCGAGACCTCATATCTGGCTTCTACGGTG 3292  
Qy ATCCCAATGAGGGAATGTGAGCAATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3133  
Db ATCCCAATGAGGGAATGTGAGCAATGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 3352  
Qy ACACCAACAGGCGAGTGTCTGACACCAAGGCGCTCAAGGGGAGGAGTGCACG 3193  
Db ACACCAACAGGCGAGTGTCTGACACCAAGGCGCTCAAGGGGAGGAGTGCACG 3412  
Qy TATGTGAGGTAGAAAATCGATACCAAGGAACCTCTCTAGAGGAACATGTTATTACTC 3253  
Db TATGTGAGGTAGAAAATCGATACCAAGGAACCTCTCTAGAGGAACATGTTATTACTC 3472  
Qy TTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTACAG 3313  
Db TTCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTACAG 3532  
Qy CTATCAATTTTGGCTACTCTGAGCAACAAACAGGGATTTGGACATGTTTCATCAATG 3373  
Db CTATCAATTTTGGCTACTCTGAGCAACAAACAGGGATTTGGACATGTTTCATCAATG 3592  
Qy CCTCCAAAGATTTCAACCTCAACATCAGTGGGCTGCGAGTTTCTCAGTGGAAACCCAG 3433  
Db CCTCCAAAGATTTCAACCTCAACATCAGTGGGCTGCGAGTTTCTCAGTGGAAACCCAG 3652  
Qy CTGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTCAAGATGTTTCT 3493  
Db CTGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGGAGTCAAGATGTTTCT 3712  
Qy CTAATGAGAGTTGATTTTTCGCAACCAACCAATATCACTTTCTTTTGTATTGTGAGTA 3553  
Db CTAATGAGAGTTGATTTTTCGCAACCAACCAATATCACTTTCTTTTGTATTGTGAGTA 3772  
Qy ATTTCACTGGCCCATCAAAATTCAGGT 3581  
Db ATTTCACTGGCCCATCAAAATTCAGAT 3800

## RESULT 3

US-09-245-041-1  
; Sequence 1, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8827  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-1

Query Match 79.1%; Score 2845.2; DB 3; Length 8827;  
Best Local Similarity 90.3%; Pred. No. 0;  
Matches 3064; Conservative 0; Mismatches 328; Indels 2; Gaps 2;  
Qy 189 ACTAACTGGATCTTCTCGGTTTGTGACAGATGGACCTGGAATTTATAAATCAAAACGAA 248  
Db 677 ACTAACTGGCTCTTCTCGATTTGTAACAGATGGACCTGGAATTTATAAATCAAAACGAA 736  
Qy 249 GTGACAGTGGCTCATTTGAGGACAGCCAAATAGATATATGAGCTTCTGCTTCAATCATTT 308  
Db 737 GTGACATGGCTCATTTGAGGACAGCCAAATAGATATATGAGCTTCTGCTTCAACCATTT 796  
Qy 309 TGCTACAGAGTGTAGTTTGGGACCAATTTATATGTTTATCATGGGACTCAATTTATGCACC 368  
Db 797 TGCTACAGATGTAGCTGGACCAATTTATATGTTTATGATGGGACTCAATCTACGCACC 856  
Qy 369 GGTAGTCTGCTCATTTAGTGGCTCATTTGTTCTGAGAGATGGAATGAGACTGTCTCC 428  
Db 857 TGTGATTTGCTGCTTTAGTGGCTCATTTGTTCTGAGAGATGGAATGAGACTGTCTCC 916  
Qy 429 TGAGGTTGTTGCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 488  
Db 917 TGAGGTTGTTGCCACATCAGGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 976  
Qy 489 TTTGCTGCTGATTTATATATCTTACAGTTTGTGATATGTTGCTGCTGCTGCTGCTGCTGCT 548  
Db 977 TCTGCTGCTGATTTATATATCTTACATTTTGTGATATGTTGCTGCTGCTGCTGCTGCTGCT 1036  
Qy 549 AGGAGAGTGAAGATCAGTATAGCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 608  
Db 1037 AGGAGAGTGAAGATCAGTATAGCAGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1096  
Qy 609 GAAAGTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 668  
Db 1097 GAAAGTGAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1156  
Qy 669 CATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 728  
Db 1157 CATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1216  
Qy 729 ATGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 788  
Db 1217 ATGTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1276  
Qy 789 AAAGCTCCAGAGCATCTCATAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 848  
Db 1277 AAAGCTCCAGAGCATCTCATAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1336  
Qy 849 AGGATATATGTTCAACCACTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 908  
Db 1337 CGGATATATGTTCAACCACTCAGATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1396  
Qy 909 GGAGTGGCTTCCACTAAACCGTTCTGTAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 968  
Db 1397 GGAATGGCTTCCACTAAACCGTTCTGTAACAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1456  
Qy 969 GGCATTATACAGATTAATTTACATGATGAGGAAATTTGATTAATGATTAATGATTAATGAT 1028  
Db 1457 GGCATTATACAGATTAATTTACATGATGAGGAAATTTGATTAATGATTAATGATTAATGAT 1516  
Qy 1029 GACCAATGAGTGTGAGAGTTTTCATTAATGAGTCAATGAGTGGTGGTGGTGGTGGTGGTGGT 1088  
Db 1517 GACCAATGAGTGTGAGAGTTTTCATTAATGAGTCAATGAGTGGTGGTGGTGGTGGTGGTGGT 1576  
Qy 1089 GGCAGAGGAGGAGTGTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1148  
Db 1577 AGCTAAGGATCAGTATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1636  
Qy 1149 CCGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1208  
Db 1637 CCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1696

QY 1209 GCAGGAATATGATTGGATAGGAACAACATGGAGTATATTAACACACCCAGGGTGCCTTGT 1268  
DB 1697 GCAGGAATATGACTTGGAAAGAAACAACATGGAGTATATTAACATATCTCAGGGTGTCTTGT 1756  
QY 1269 GCAAGGGGGTTACGGCCATAGCAGTGTGTACGACCATAGACACAGAGGGCCCTATAGTTCA 1328  
DB 1757 GCAAGGGGGTTATGGCCACAGTGTGTGTATGATGACAGCAAGAGCTCTGTACGTCA 1816  
QY 1329 TGTGGCTACAAGGCTTTCAAGTCCCAATAGTACCGGCTTGACAGTGTATCTTACCGATA 1388  
DB 1817 TGTGGCTACAAGGCTTTCAAGTCCCAATAGTACCGGCTTGACAGTGTATCTACAGATA 1876  
QY 1389 TGATGTGGATACCCAGATGTGACATCTTTAAGGACAGCGGATTTTTCGTACTTGCA 1448  
DB 1877 CGATGTGGATCTCAGATGTGACCAATCTTTAAGGACAGCGGATTTTTCGTACTTGCA 1936  
QY 1449 CACAGCTGTGATAGTGTGAAACCATGCTGTGTGTGTGGGGGAAACACACAAATGACAC 1508  
DB 1937 TACAGCTGTGATAGTGTGAAACCATGCTGTGTGTGTGGGGGAAACACACAAATGACAC 1996  
QY 1509 ATCTATGAGCCATGGCCCAAAATGCTTCTCTCAGATTTTCATGGCCTATGACATTCGCTG 1568  
DB 1997 TTCCATGAGCCACGGTGCACAAATGCTTCTCTCGGACTTCAATGGCTTATGACATTCGCTG 2056  
QY 1569 TGACCGTGTGATGTGCTTCCAGACCTGA-TTCCACCATGATGTCAACAGATTTTGCCCA 1627  
DB 2057 TGACCGATGTGATGTGCTTCCAGACCTGAGCTCCATCATGATGTCAACAGATTTTGCCCA 2116  
QY 1628 TTCCAGCAGCTTTACACACAGCACCATGTATGTGTGCTGTGTGTCAATAGTCTCCTCC 1687  
DB 2117 TT-CAGCAGCTTTGTACACACAGCACCATGTATGTGTGCTGTGTGTCAACAGCCTCCTCC 2175  
QY 1688 TCACCGATCTCTGTGTTTACCTCGGAACAGTGTGATGTGCTGTGCTGAGTGAAGCGCTT 1747  
DB 2176 TCAGTGAAGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2235  
QY 1748 GTTTAGCAGCAGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1807  
DB 2236 GTGTGGCAGCAGGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2295  
QY 1808 GTGTGGCGCTGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1867  
DB 2296 CCTGGAGTGTGGCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2355  
QY 1868 GAACCTTTGACCATGACAGATGTGACAGCACAAGATTTTAAATCAGAAATGTGTGTGTGTGT 1927  
DB 2356 GAACCTTTGACCATGACAGATGTGACAGCACAAGATTTTAAATCAGAAATGTGTGTGTGTGT 2415  
QY 1928 CCAATGACTGCCACTGGTGTGCAATGACCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1987  
DB 2416 CCAATGACTGCCACTGGTGTGCAATGACCAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2475  
QY 1988 GCCAGATCTCAATTTTAGTGTATGAAATTTGCCCAAGGATACCTATGTACTACTGTA 2047  
DB 2476 GCCAGATCTCAATTTTAGTGTATGAAATTTGCCCAAGGATACCTATGTACTACTGTA 2535  
QY 2048 ACAAGAAAGACAGCTGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2107  
DB 2536 ATAAGAAACACAGCTGACGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2595  
QY 2108 ATCAGGAGTGTATGCTTCCCGGAAATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2167  
DB 2596 ATCAGGAGTGTATGCTTCCCGGAAATATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2655  
QY 2168 ACTCATGTGTGAAATTTACTTCCCAAGGAGATTTATGACAAATGTGTGTGTGTGTGTGTGTGTGT 2227  
DB 2656 ACTCATGTGTGAAATTTACTTCCCAAGGAGATTTATGACAAATGTGTGTGTGTGTGTGTGTGTGT 2715  
QY 2228 GGAACCAATATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2287  
DB 2716 GGAACCAATATGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2775  
QY 2288 AGCAGCTGCGAATTAATGACAGTCTATCTCAGAGCATGTCTCAAGCTCACTTAAACCCCATGGG 2347

DB 2776 AGCAGCTTCCGATTAATGCAATCATCTCAAGTATGTCAAGCTCACTCTGACTCCATGGG 2835  
QY 2348 TCGGCTTTCCGAAGATCAATGTGTCTTACTGTCTGCTGGAGATATATGCTCCCAATTTTACAA 2407  
DB 2836 TTGTCTTTCCGAAGATCAATGTGTCTTACTGTCTGCTGGAGATATATGCTCCCAATTTTACAA 2895  
QY 2408 ATAGTTTACTACAGTGTGATCCCTGTGAGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2467  
DB 2896 ATAGTTTGTCTGAGTGTGATCCCTGTGAGCCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2955  
QY 2468 CAGAACCCAGTGTGAGGACTGAAGCTGTCAACCTGCATCAACCCACTCAATGTGTGTGTGTGT 2527  
DB 2956 CAGAACCCAGTGTGAGGACTGAAGCTGTCAACCTGCATCAACCCACTCAATGTGTGTGTGTGT 3015  
QY 2528 TCTGTGAAAGGCTGTGAAACCCAGTGTGAGGACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2587  
DB 3016 TCTGTGAAAGGCTGTGAAACCCAGTGTGAGGACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3075  
QY 2588 CAGCATGTGAGATTTGACACGAGGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2647  
DB 3076 CAGCATGTGAGATTTGACACGAGGAGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 3135  
QY 2648 AGTGTGTGAGTCTCAATGTGCTTATGTGCTCTTCCCTTTTGGCCAGTGTATGGAATGGT 2707  
DB 3136 AGTGTGTGAGTCTCAATGTGCTTATGTGCTCTTCCCTTTTGGCCAGTGTATGGAATGGT 3195  
QY 2708 ATAGATGAGACCTGTGCTTCCCTTCAAAATTTGTTTGGCTACTGTACCTGTGTGTGTGTGTGT 2767  
DB 3196 ATAGATGAGACCTGTGCTTCCCTTCAAAATTTGTTTGGCTACTGTACCTGTGTGTGTGTGTGT 3255  
QY 2768 TGGAGCAACAGGCTGTGCTTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2827  
DB 3256 TGGAGCAACAGGCTGTGCTTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3315  
QY 2828 TAGAGGTTTCTTAAGAGGACAGTGTGAGTGTGCTTGGAGCCCTTACAGGAAATTTCT 2887  
DB 3316 TAGAGGTTTCTTAAGAGGACAGTGTGAGTGTGCTTGGAGCCCTTACAGGAAATTTCT 3375  
QY 2888 ATCCACAGCCCTGTCTCAATTTCCAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2947  
DB 3376 ATCCACAGCCCTGTCTCAATTTCCAGCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3435  
QY 2948 TCATTTACTGTCTCAGTGTGCAATGTCAACCGGCTTACAGTGTGTGTGTGTGTGTGTGTGTGTGT 3007  
DB 3436 TCATTTACTGTCTCAGTGTGCAATGTCAACCGGCTTACAGTGTGTGTGTGTGTGTGTGTGTGTGT 3495  
QY 3008 GTGAGAGTGTGAGGACTGTGAGGACTGTGAGGACTGTGAGGACTGTGAGGACTGTGAGGACTGT 3067  
DB 3496 GTGAGAGTGTGAGGACTGTGAGGACTGTGAGGACTGTGAGGACTGTGAGGACTGTGAGGACTGT 3555  
QY 3068 AGGTTGATTTCCCAATGTGAGGAGGAAATGTGAGGAGGAAATGTGAGGAGGAAATGTGAGGAGG 3127  
DB 3556 AGGTTGATTTCCCAATGTGAGGAGGAAATGTGAGGAGGAAATGTGAGGAGGAAATGTGAGGAGG 3615  
QY 3128 TGTGCAACACCAACACCGGCTGTCTTCTGCAACCAAGGCTTCAAGGGGACCGAGT 3187  
DB 3616 TGTGCAACACCAACACCGGCTGTCTTCTGCAACCAAGGCTTCAAGGGGACCGAGT 3675  
QY 3188 GCCAGCTATGTGAGGAGGAAATGTGAGGAGGAAATGTGAGGAGGAAATGTGAGGAGGAAATGTG 3247  
DB 3676 GCCAGCTATGTGAGGAGGAAATGTGAGGAGGAAATGTGAGGAGGAAATGTGAGGAGGAAATGTG 3735  
QY 3248 ATACTCTTTTATTGACTATCAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3307  
DB 3736 ATACTCTTTTATTGACTATCAGTGTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3795  
QY 3308 ACACAGCTATCAATTTTGTGCTACTCTGTGACGAAACAAACAGGGATTTTGACATGTTTCA 3367  
DB 3796 ACACAGCTATCAATTTTGTGCTACTCTGTGACGAAACAAACAGGGATTTTGACATGTTTCA 3855  
QY 3368 TCAATGCTTCCAGAAATTTTCAACCTCAACATCACTGCTGGCTGCGAGTCTTCTCAGCTGGA 3427

Db	3856	TC	AATGCTCC	AAAACTT	CAACCTCA	CAATCA	CTGGGCCA	CCAGCTT	CC	CAGCGGAA	3915	
Qy	3428	CC	CAGGCTGG	AGAGAGAT	GCTCTT	CTTTCA	AAACCA	CAATTA	AGGAGT	TACAAGATA	3487	
Db	3916	CC	CAGACTGG	AGAGAGTG	CTGTTT	TTCAAAA	CCAA	CACTC	AGGAAT	TACAAGATA	3975	
Qy	3488	GT	TCTCTAAT	GAGAAGTT	TGATTTT	CGCAAC	CA	CCCAATAT	CACATTT	CTTTGTTTATG	3547	
Db	3976	GC	TCTCTAAT	GAGAAATTT	TGATTTT	CGCAAC	CA	CCATC	CAACAT	CACATTT	CTTTGTTTATG	4035
Qy	3548	TC	AGTAAATTT	CACCTGG	CCCATCA	AAATTC	AGGT				3581	
Db	4036	TC	AGTAAATTT	CACCTGG	CCCATCA	AAATTC	AGAT				4069	

```

RESULT 4
US-09-245-041-18
; Sequence 18, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-03-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: PASTSEQ for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-18

```

Query Match	56.43;	Score	2028.8;	DB	3;	Length	2625;
Best Local Similarity	97.04;	Pred. No.	0;				
Matches	2088;	Conservative	0;	Mismatches	62;	Indels	2;
Gaps	2;						
Qy	15	GGCGGCAACTGAGCAAGGCTGAGGAGGAGGAGACGGCGGCGACGGCAGCGCTCGCGGGCAG	74				
Db	234	GGAGGCGAGGCGCGCTCGGTGGCGGGCGGGGTGTCGGGTCGGCGCAGCCGAGGCGAA	293				
Qy	75	GAGGGGGGGCGGCACTGTCTCAACGGCGGTTCGTGCAACCGCTGGCACGGCGCAGTGCCT	134				
Db	294	GGAAATGTGACCGGCCGTGTGTCAACGGCGGTTCGTGCAACCGCTGGCACCGCGCAGTGCCT	353				
Qy	135	CTGCCCGCGCGGTGGTGGCGGAGCAATGCCAGCACTGCGGGGGCGCGCTTCAGACTPAAC	194				
Db	354	CTGCCCGCGCGGTGGTGGCGGAGCAATGCCAGCACTGCGGGGGCGCGCTTCAGACTPAAC	413				
Qy	195	TGGATCTTCTGGGTTTGTGCACAGATGCACCTGGAAATTTATAATACAAAAACGAAGTGCAC	254				
Db	414	TGGATCTTCTGGGTTTGTGCACAGATGCACCTGGAAATTTATAATACAAAAACGAAGTGCAC	473				
Qy	255	GTGCTCATTTGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATCAATTTGCTAC	314				
Db	474	GTGCTCATTTGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATCAATTTGCTAC	533				
Qy	315	AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT	374				
Db	534	AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGGACTCAATTTATGCACCGCTAGT	593				
Qy	375	TGCTGCATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGT	434				
Db	594	TGCTGCATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGT	653				
Qy	435	TGTTGCCACATCAGGTTATGCTTTCGTCGCAATTTTTTTTATGATGCTGCTCTATAAATTGAC	494				

Db	654	TGTTGCACATCAGGTTATGCGCTTGTGTCAGATTTTTTTTAGTGAAGCTGCTTATATAATTTGAC	713
Qy	495	TGGAATTTAAATATTACTTTACAGTTTTCATATGTGTCCAAATAAATGCTCAGGCCGAGGAGA	554
Db	714	TGGATTTAAATATTACTTTACAGTTTTCATATGTGTCCAAATAAATGCTCAGGCCGAGGAGA	773
Qy	555	GTGTAAGATCAGTAAATAGCAGCGATCTGTTGAAATGTGAATGTCTGAAAACTGGAAGG	614
Db	774	GTGTAAAGATCAGTAAATAGCAGCGATCTGTTGAAATGTGAATGTCTGAAAACTGGAAGG	833
Qy	615	TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCATCTG	674
Db	834	TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCATCTG	893
Qy	675	CAATTCAGGTGATGTGACAGGATGCTCCTGCTCTCAGACTGGCAGGGTCTCGATGTTTC	734
Db	894	CAATTCAGGTGATGTGACAGGATGCTCCTGCTCTCAGACTGGCAGGGTCTCGATGTTTC	953
Qy	735	AGTTCCTGTACCAAGCTAACCCAGTCATTTTCGACTCGAGAGGAATATTCTAACTTTAAAGCT	794
Db	954	AGTTCCTGTACCAAGCTAACCCAGTCATTTTCGACTCGAGAGGAATATTCTAACTTTAAAGCT	1013
Qy	795	CCCAGAGCATCTCATAAAGCTGTGTCATGTGAAACATTTATGTGGTGTGTTGGAGGATA	854
Db	1014	CCCAGAGCATCTCATAAAGCTGTGTCATGTGAAACATTTATGTGGTGTGTTGGAGGATA	1073
Qy	855	TATGTTCAACCACTCAGATTTAACATGTGTTCTAGCGTATGACCTTCCTCTAGGAGTG	914
Db	1074	TATGTTCAACCACTCAGATTTAACATGTGTTCTAGCGTATGACCTTCCTCTAGGAGTG	1133
Qy	915	GCITCCACTAAACCGTTTCTGTGAACATGTGGTTGTGTAGATATGGTCAITCTTTGGCATT	974
Db	1134	GCITCCACTAAACCGTTTCTGTGAACATGTGGTTGTGTAGATATGGTCAITCTTTGGCATT	1193
Qy	975	ATACAGGATAAATTTACATGTATGAGAGAAATTTGATTTCACTCGGGAATGTGACCAA	1034
Db	1194	ATACAGGATAAATTTACATGTATGAGAGAAATTTGATTTCACTCGGGAATGTGACCAA	1253
Qy	1035	TGAGTTGAGAGTTTTTTCACATTCATTAATGAGTCATGGGTGTGTGTACCCCTTAAGGCAAA	1094
Db	1254	TGAGTTGAGAGTTTTTTCACATTCATTAATGAGTCATGGGTGTGTGTACCCCTTAAGGCAAA	1313
Qy	1095	GGAGCATATGCGATGTTGGGCACTCTGCACACATTTGTTTACACTGAAGAAATGCGCGAGT	1154
Db	1314	GGAGCATATGCGATGTTGGGCACTCTGCACACATTTGTTTACACTGAAGAAATGCGCGAGT	1373
Qy	1155	GGTCATGCTCGTCATCTTTGGTCAGTCCCTCTATGGATATATAAGCAATGTGCAGGA	1214
Db	1374	GGTCATGCTCGTCATCTTTGGTCAGTCCCTCTCTATGGATATATAAGCAATGTGCAGGA	1433
Qy	1215	ATATGATTTGGATAAGAACACATGGAGTATATTACACCCAGGGTCCCTTTGCGAAGG	1274
Db	1434	ATATGATTTGGATAAGAACACATGGAGTATATTACACCCAGGGTCCCTTTGCGAAGG	1493
Qy	1275	GGGTTACGGCCATAGCAGTCTTTACGACCATAGACACAGGGCCCTATACGTTTCATGTGG	1334
Db	1494	GGGTTACGGCCATAGCAGTCTTTACGACCATAGACACAGGGCCCTATACGTTTCATGTGG	1553
Qy	1335	CTCAAGGCTTTTCAGTGCCCAATTAAGTACCGGCTTTCAGATGATCTCTACCGATATGATGT	1394
Db	1554	CTCAAGGCTTTTCAGTGCCCAATTAAGTACCGGCTTTCAGATGATCTCTACCGATATGATGT	1613
Qy	1395	GGATACCCAGATGTGGACCAATCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGC	1454
Db	1614	GGATACCCAGATGTGGACCAATCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGC	1673
Qy	1455	TGTGATAGTCAGTGGAAACCATGCTGGTGTGTTGGGGGAAAACACACAATGACACATCTAT	1514
Db	1674	TGTGATAGTCAGTGGAAACCATGCTGGTGTGTTGGGGGAAAACACACAATGACACATCTAT	1733
Qy	1515	GAGCCATGGCGCAAAATGCTTCTTTCAGATTTTCATGGCCTATAGCAATTCGCTGTGACCG	1574
Db	1734	GAGCCATGGCGCAAAATGCTTCTTTCAGATTTTCATGGCCTATAGCAATTCGCTGTGACCG	1793

QY 1575 CTGCTCAGTCTCTCCAGACCTGAT-TCACCAATGATGTCAACAGAAATTTGGCCATTCCAG 1633  
Db 1794 CTGCTCAGTCTCTCCAGACCTGATCTCCACCAATGATGTCAACAGAAATTTGGCCATT-CAG 1852  
QY 1634 CAGTCTTACACACAGACACCATGTATGTGTTCGTTGTTTCAATAGTCTCTCCCTCAGCG 1693  
Db 1853 CAGTCTTACACACAGACACCATGTATGTGTTCGTTGTTTCAATAGTCTCTCCCTCAGCG 1912  
QY 1694 ACATCTCTGTTATTCACCTCGGAACAGTGTGATCGCATCGAGTGAAGCCGCTGTTTAG 1753  
Db 1913 ACATCTCTGTTATTCACCTCGGAACAGTGTGATCGCATCGAGTGAAGCCGCTGTTTAG 1972  
QY 1754 CAGCAGACCTGTTATTCGTTGTTGTGTGGAACACAGGGTCTCTCAGTGTATCTCTGTGG 1813  
Db 1973 CAGCAGACCTGTTATTCGTTGTTGTGTGGAACACAGGGTCTCTCAGTGTATCTCTGTGG 2032  
QY 1814 CGTCTGCAACTGTATGAACAGAAAGAAAGTTAAATCAGAAATGTTTTCCAAAAGAACTC 1873  
Db 2033 CGTCTGCAACTGTATGAACAGAAAGAAAGTTAAATCAGAAATGTTTTCCAAAAGAACTC 2092  
QY 1874 TTGACCATGACAGATGTGACACACACAGATTTGTTACAGTCTCCAGACCAACCAATG 1933  
Db 2093 TTGACCATGACAGATGTGACACACACAGATTTGTTACAGTCTCCAGACCAACCAATG 2152  
QY 1934 ACTGCCACTGTGCAATGACCAATGTGTCTCCAGGAACCAACAGTCTCAGAAAGCCAGA 1993  
Db 2153 ACTGCCACTGTGCAATGACCAATGTGTCTCCAGGAACCAACAGTCTCAGAAAGCCAGA 2212  
QY 1994 TCTCCATTTTATGATATGAATGCCCCAAGGATAACCCATGTATCTACTGTAAACAGA 2053  
Db 2213 TCTCCATTTTATGATATGAATGCCCCAAGGATAACCCATGTATCTACTGTAAACAGA 2272  
QY 2054 AGACCACTGACAGAGCTGTGCTCCAGACAGAACTGCCAGTGGGAGCCCCCGGAATCAGG 2113  
Db 2273 AGACCACTGACAGAGCTGTGCTCCAGACAGAACTGCCAGTGGGAGCCCCCGGAATCAGG 2332  
QY 2114 AGTGCATTTGCCCTCCCGAAATATCTGTGGCAATGCTGGCAATTTGGTTGG 2165  
Db 2333 AGTGCATTTGCCCTCCCGGTTAGGCTTTGACGGTCTATCTGTGTGTGTGG 2384

## RESULT 5

US-09-245-041-8  
; Sequence 8, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-8

Query Match 39.7%; Score 1428.8; DB 3; Length 2419;  
Best Local Similarity 89.4%; Pred. No. 0;  
Matches 1560; Conservative 0; Mismatches 182; Indels 2; Gaps 2;

QY 189 ACTAACTGATCTCTCTGTTTGTGACAGATGACCTGGAAATATATAATACAAAACGAA 248  
Db 677 ACTAACTGCTCTCTCGATTTGTAAACAGATGACCTGGGAATTTATAAATATAAGACGAA 736

QY 249 GTGCACGTGGCTCATTTGAAGACAGCCAAATAGAAATATAGAGACTTCGTTTCAATCATTT 308  
Db 737 GTGCACATGGCTCATTTGAAGACAGCCAAATAGAAATATAGAGACTTCGTTTCAACCATTT 796  
QY 309 TGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCACC 368  
Db 797 TGCTACAGAAATGAGTGGGACCAATTTATATGTTTATGATGGGACTCAATTTACGCACC 856  
QY 369 GCTAGTTGCTGCATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCC 428  
Db 857 TCTGATTTGCTGCCCTTTAGTGGCTCATTTGTTCTGAAAGAGATGGCAATGAGAGCGGTCC 916  
QY 429 TGAGTGTGTTGCCACATCAGGTTATGCTTGTCTGTCATTTTTTTAGTGTGCTGCTTATAA 488  
Db 917 TGAGTCTACTGTCACTTCAGGTTATGACTGTGTCATTTTTTTCAGTGTGCTGCTTATAA 976  
QY 489 TTTGACTGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 548  
Db 977 TCTGACTGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1036  
QY 549 AGGAGAGTGTAAAGTCAAGTATAGCAGGATACCTGTTGAATGTGAATGTTCTGAAACTG 608  
Db 1037 AGGAGAGTGTAAAGTCAAGTATAGCAGGATACCTGTTGAATGTGAATGTTCTGAAACTG 1096  
QY 609 GAAAGGTGAAGCATGTGACATTTCTACTGTACAGCAACTGTGGTTTTCTCATCGAGG 668  
Db 1097 GAAAGGGAGTGTGTGACATTTCTACTGTACAGCAACTGTGGTTTTCTCATCGAGG 1156  
QY 669 CATCTGCAATTTCAAGTGTGACAGAGTGTCTGCTTCTCAGACTGCGAGGGTCTCTGG 728  
Db 1157 CATCTGTAATGACAGGATACAGAGGGTGTCTCTGCTTTCTCAGTGGAGGGTCTCTGG 1216  
QY 729 ATGTTTCACTTCTGTACAGCTAACCACTCATTTTGGACTCGAGAGAAATTTCTAACTT 788  
Db 1217 ATGTTTCAATTTCTGTCCAGCTAACCACTCATTTTGGACTCGAGAGAAATTTCTGATTT 1276  
QY 789 AAAGTCTCCACAGAGATCTCATAAAGCTGTGGTCAATGGAAACATTAATGTGGTGTGG 848  
Db 1277 AAAGCTTCCACAGAGCTCTCAATAAGCTGTGGTCAATGGAAATTAATAATGTGGTGTGG 1336  
QY 849 AGGATATATGTTCAACCACTCAGATTAATAATGTTTCTAGCGTATGACCTTGTCTCTAG 908  
Db 1337 CGGATATATGTTCAACCACTCAGATTAATAATGTTTCTAGCGTATGACCTTGTCTCTAG 1396  
QY 909 GGAGTGGCTTCCACTAAACCGTTCTGTGAACAAATGTGGTGTGTAGATATGGTCTCTTT 968  
Db 1397 GGAATGGCTTCCACTAAACCAATTTCTGTGAACAGTGTGGTGTGTAGATATGGTCTCTTT 1456  
QY 969 GGCATTATACAGGNTAAATTTTACATGTTAGGAGGAAATTTGATTTCACTGGGAATGT 1028  
Db 1457 GGCATTATACAGGNTAAATTTTACATGTTAGGAGGAAATTTGATTTCAACAGGGAATGT 1516  
QY 1029 GACCAATGAGTTGAGAGTTTTTCACTTCAATTAATGAGTGTGGTGTGTGTGACCCCTAA 1088  
Db 1517 GACCAATGAGTGTGAGAGTTTTTCACTTCAATTAATGAGTGTGGTGTGTGTGACCCCTAA 1576  
QY 1089 GGCAGAGAGAGTATGAGTGTGGGCTGTGACACATTTGTTACACTGAGAGATGG 1148  
Db 1577 AGCTAAGGATCAGTATGAGTGTGGGCTGTGACACATTTGTTACACTGAGATCTGG 1636  
QY 1149 CCGAGTGGTCTATGCTGGTCTATTTGTTGCTACTGCTCTCTATGGATATATAAGCAATGT 1208  
Db 1637 CCGTGTGGTCTATGCTGGTCTATTTGTTGCTACTGCTCTCTATGGATATATAAGCAATGT 1696  
QY 1209 GCAGGAATATGATTTGATTAAGACACATGAGTATATTAACACCCAGGGTGTCTTGT 1268  
Db 1697 GCAGGAATATGATTTGATTAAGACACATGAGTATATTAACACTCAGGGTGTCTTGT 1756  
QY 1269 GCAAGGGGGTTACGGCCATAGCAGTGTGTTACGACATAGGACCAAGGGCCCTATACGTTCA 1328  
Db 1757 GCAAGGGGGTTATGGCCACAGTAGTGTGTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816









```
Db 186 CTTGGAAGATCAATGTGTCCTACTGCTGGTGGGAAGATATGTCCCATTTTCAAAATAGT 127
QY 2413 TTACTACAGTGGATGCGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAA 2472
    |||||
Db 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAA 67
QY 2473 CCAGTACTCGGGACTGAAGGTGCAACTGATCAACCCACTCAATGTAGTGTCTGT 2532
    |||||
Db 66 CCAGTACTCGGGACTGAAGGTGCAACTGATCAACCCACTCAATGTAGTGTCTGT 7
QY 2533 GAAAGG 2538
    |||||
Db 6 GAAAGG 1

RESULT 9
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; CURRENT APPLICATION NUMBER: US/09/736,457
; CURRENT FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 1864
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Query Match 6.8%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.1e-64;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 CTGCGAATAATGAGTCAATGTGCTCTACTGCTGGGAAGATATGTCCCATTTTCAAAATAGT 2412
    |||||
Db 246 CTGCGAATAATGAGTCAATGTGCTCTACTGCTGGGAAGATATGTCCCATTTTCAAAATAGT 187
QY 2353 CTTGCGAAGATCAATGTGCTCTACTGCTGGGAAGATATGTCCCATTTTCAAAATAGT 2412
    |||||
Db 186 CTTGCGAAGATCAATGTGCTCTACTGCTGGGAAGATATGTCCCATTTTCAAAATAGT 127
QY 2413 TTACTACAGTGGATGCGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAA 2472
    |||||
Db 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAA 67
QY 2473 CCAGTACTCGGGACTGAAGGTGCAACTGATCAACCCACTCAATGTAGTGTCTGT 2532
    |||||
Db 66 CCAGTACTCGGGACTGAAGGTGCAACTGATCAACCCACTCAATGTAGTGTCTGT 7
QY 2533 GAAAGG 2538
    |||||
Db 6 GAAAGG 1

RESULT 10
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966
```

```
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1688
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Query Match 6.8%; Score 246; DB 4; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.1e-64;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 CTGCGAATAATGAGTCAATGTGCTCTACTGCTGGGAAGATATGTCCCATTTTCAAAATAGT 2352
    |||||
Db 246 CTGCGAATAATGAGTCAATGTGCTCTACTGCTGGGAAGATATGTCCCATTTTCAAAATAGT 187
QY 2353 CTTGCGAAGATCAATGTGCTCTACTGCTGGGAAGATATGTCCCATTTTCAAAATAGT 2412
    |||||
Db 186 CTTGCGAAGATCAATGTGCTCTACTGCTGGGAAGATATGTCCCATTTTCAAAATAGT 127
QY 2413 TTACTACAGTGGATGCGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAA 2472
    |||||
Db 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAA 67
QY 2473 CCAGTACTCGGGACTGAAGGTGCAACTGATCAACCCACTCAATGTAGTGTCTGT 2532
    |||||
Db 66 CCAGTACTCGGGACTGAAGGTGCAACTGATCAACCCACTCAATGTAGTGTCTGT 7
QY 2533 GAAAGG 2538
    |||||
Db 6 GAAAGG 1

RESULT 11
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-966
```

Query Match 6.8%; Score 246; DB 4; Length 246;  
Best Local Similarity 100.0%; Pred. No. 2.1e-64; Indels 0; Gaps 0;  
Matches 246; Conservative 0; Mismatches 0;

QY 2293 CTGCGAATAATGCACTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGC 2352  
DB 246 CTGCGAATAATGCACTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGC 187

QY 2353 CTTCGGAAGATCAATGTCTCTACTGTCTGGGAGATATGTCCTTACAAATAGT 2412  
DB 186 CTTCGGAAGATCAATGTCTCTACTGTCTGGGAGATATGTCCTTACAAATAGT 127

QY 2413 TTACTACAGTGGATGCCGCTGTGAGCCAGTGTGCTGGAATTTTATCAGAA 2472  
DB 126 TTACTACAGTGGATGCCGCTGTGAGCCAGTGTGCTGGAATTTTATCAGAA 67

QY 2473 CCAGTACTCGGGAGTGAAGGCTGCAACCTGTGATCAACCCACTCAATGTAGTCTGT 2532  
DB 66 CCAGTACTCGGGAGTGAAGGCTGCAACCTGTGATCAACCCACTCAATGTAGTCTGT 7

QY 2533 GAAAGG 2538  
DB 6 GAAAGG 1

RESULT 12  
US-09-833-381-1916  
; Sequence 1916, Application US/09833381  
; Patent No. 6672186  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; PRIOR FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1916  
; LENGTH: 415  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(415)  
; OTHER INFORMATION: n = A,T,C or G

US-09-833-381-1916

Query Match 5.9%; Score 212.8; DB 4; Length 415;  
Best Local Similarity 73.6%; Pred. No. 3.7e-54; Indels 0; Gaps 0;  
Matches 271; Conservative 0; Mismatches 97;

QY 200 CTCTCGGTTTGTGACAGATGGACCTGGAAATTAATAACAAAGTGCACGTGGC 259  
DB 5 CGTCCGATATTTAAGATGGCCCAATTAACATAATAAATTAATGTTACTTGGC 64

QY 260 TCATTGAAGACCCAAATAGATATGAGACTTGTTCATCATTTTGTACAGAT 319  
DB 65 TCATTGAAGGCTATCCAAATGCACTGTAAAGATTAAGATTCATCATTTGTACAGAAT 124

QY 320 GTAGTGGGACATTTATATGTTTATGATGGGACTCAATTTATGACCCCTAGTGTCTG 379  
DB 125 GTAGTGGGACATATGTTTATGATGGGACTCAATTTATGACCCCTTTAATAGCTG 184

QY 380 CATTTAGTGGCTCATTTGTTCTGAGAGAGATGCAATGCTCCCTGAGTTTGTG 439  
DB 185 TACTTAGTGGTTTGATAGTCCCTGAAATGAGGCAATGCTGCTGAAATTTGTTA 244

QY 440 CCACATGAGTTATGCTGTGCTATTTTATGATGCTGTTTAAATTTGACCTGGAT 499  
DB 245 CTACATCTGGCTATGCACTGTTACATTTTATGATGCTGGTATATCTAATCTAATCTGGTT 304

QY 500 TTAATATTACTTACAGTTTTCATATGTTCCAAATAACTGCTCAGGCCGAGGAGTGTGA 559  
DB 305 TCACATTTTCTATTCAATCAATTTCTTGTCTCAACAATTTCTTGGTCAATGGGAAGTGA 364

QY 560 AGATCAGT 567  
DB 365 CAACTAGT 372

RESULT 13  
US-09-245-041-5  
; Sequence 5, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER FILING DATE: 1998-07-21  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 90050  
; TYPE: DNA  
; ORGANISM: Mus musculus  
; US-09-245-041-5

Query Match 5.6%; Score 202.8; DB 3; Length 90050;  
Best Local Similarity 81.8%; Pred. No. 2.2e-49;  
Matches 234; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY 947 TGTGTAGATGTCATCTTTGGCATATACAGGATAAAATTTACATGATGAGGAA 1006  
DB 6931 TGGTTTAAAAATCCTGTTTACATCTTTTATAGGATAAATCTACATGATGAGGAA 6990

QY 1007 AAATTGATTCACCTGGGAATGTGACCAATGATGTTTTCAGAGTTTTCACATTCATATGAT 1056  
DB 6991 AAATTGATTCACAGGGAACGTGACCAATGATGAGTATTCATATTCATAATGAAT 7050

QY 1067 CATGGTGTGTCACCCCTAAGGCAAGGAGCATGATGAGTGTGGGCACTCTGCAC 1126  
DB 7051 CATGGTATTTGTTAACTCCGAAAGCTAAGGATCATGATGAGTGTGGGCACTCTGCAC 7110

QY 1127 ACATTGTTTACACTGAAGAATGCCGAGTGGTGCATGCTGGTGCATCTTTGGTCACTGCCCTC 1186  
DB 7111 ACATTGTTTACACTGCGCTGCGTGGTGCATGTTGGTGCATCTTGGTCACTTGGCCAC 7170

QY 1187 TCTATGGATATATAAGCAATGTGAGGAATATGATTTGGATAGAA 1232  
DB 7171 TCTATGGATATATAAGCGTGTGTCAGGAATATGACTTTGGGTATGTA 7216

RESULT 14  
US-09-245-041-3/c  
; Sequence 3, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978

Search completed: March 5, 2004, 12:20:03  
Job time : 175.851 secs

; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 17056  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-3

Query Match 4.7%; Score 169.2; DB 3; Length 17056;  
Best Local Similarity 86.9%; Pred. No. 9.9e-40;  
Matches 186; Conservative 0; Mismatches 28; Indels 0; Gaps 0;  
QY 510 TTACAGTTTGTATGTCTCAATAACTGCTCAGGCCGAGGAGCTGTAAGATCAGTAA 569  
DB 4090 TTGTAGTTTGCATGTGTCGCGAATTAATGCTCAGGCCGAGGAGCTGTAAGAGCACTAA 4031  
QY 570 TACAGCGATPACTGTTGAAATGTGAATGTTCTGAAAACTGGAAGGTGAAGCATGTGACAT 629  
DB 4030 CAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACTGGAAGGTGAAGCATGTGACAT 3971  
QY 630 TCCTCACTGTACAGACAATGTGTTTCTCTCATCGAGGCATCTGCAATTCAGTGATGT 689  
DB 3970 TCCTCACTGTACAGACAATGTGTTTCTCTCATCGAGGCATCTGCAATTCAGTGATGT 3911  
QY 690 CAGAGGATGCTCCTGCTTTCTCAGACTGCGAGGCT 723  
DB 3910 CAGAGGATGCTCCTGCTTTCTCAGACTGCGAGGCT 3877

RESULT 15  
US-09-245-041-6/c  
; Sequence 6, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; FILE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 226  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-6

Query Match 3.7%; Score 132.2; DB 3; Length 226;  
Best Local Similarity 78.6%; Pred. No. 6.7e-30;  
Matches 158; Conservative 0; Mismatches 43; Indels 0; Gaps 0;  
QY 2389 GATATGTCCTCAATTCACAAATAGTTTGTGAGGCTGCAAAACACAGTGTGAAGCAGTCCCG 2448  
DB 225 GATATGTCCTCAATTCACAAATAGTTTGTGAGGCTGCAAAACACAGTGTGAAGCAGTCCCG 166  
QY 2449 GGATTCGTGGAATTTTATCAGACCCAGTACTCGGGCTGAGGCTGCAACCTGCATC 2508  
DB 165 GGCTTCGTGGGATCTTGTGAGAGCTGAGTACTCGGGGATTAAGGCTGCAACCTGCATC 106  
QY 2509 AACCCACTCAATGTAGTGTCTGTGAAGGCTGCAAAACACAGTGTGAAGCAGTCCCG 2568  
DB 105 AACCTCTCAATGGCAGCGTCTGTGAAGGCTGCAAAACACAGTGTGAAGCAGTCCCG 46  
QY 2569 ACACATGTGCTTGAGGACA 2589  
DB 45 CCAGGAGGAGCCAGACAGCA 25

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model  
Run on: March 4, 2004, 16:07:54 ; Search time 10236 Seconds  
(without alignments)  
17225.392 Million cell updates/sec

Title: US-09-787-097-11  
Perfect score: 4068  
Sequence: 1 atgtgtgcgcagcgccgc.....agcctgggacctgcacatcga 4068

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues  
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*  
29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.mam.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3988.4	98.0	8589	6	AR164816 Sequence
2	3566	87.7	4493	4	AF531101 Bos tauru
3	3556.4	87.4	3597	9	AF034957 Homo sapi
4	3499.8	86.0	4072	6	AR164817 Sequence
5	3492.4	85.9	3819	9	AF106861 Homo sapi
6	3395.4	83.5	4284	10	AB062913 Mesocric
7	3377.4	83.0	6083	10	AF116897 Mus muscu
8	3354.8	82.5	4313	10	AF116897 Mus muscu
9	3353.2	82.4	8739	10	AB038387 Rattus no
10	3326.4	81.8	8827	6	AR164807 Sequence
11	2928.8	72.0	4500	10	AB038388 Rattus no
12	2694.6	66.2	5683	10	AB038388 Mus muscu
13	2051.8	50.4	2625	6	AR164818 Sequence
14	1853	45.6	6370	6	AR164815 Sequence
15	1697	41.7	2578	9	AK000356 Homo sapi
16	1466.6	36.1	4463	9	AK127277 Homo sapi
17	1465	36.0	4140	9	AY442317 Homo sapi
18	1451.8	35.7	2419	6	AR164813 Sequence
19	1361	33.5	5632	9	AB011120 Homo sapi
20	1141.6	28.1	7302	9	AB011106 Homo sapi
21	1131	27.8	3609	6	AX503776 Sequence
22	1098	27.0	5002	10	BC027764 Mus muscu
23	817.8	20.1	1012	6	E26747 Structure a
24	704.2	17.3	4997	10	BC050020 Mus muscu
25	585.8	14.4	3297	10	BC030872 Mus muscu
26	543	13.3	543	6	AX778545 Sequence
27	463	11.4	1051	6	AR164814 Sequence
28	418.4	10.3	1818	9	BC047716 Homo sapi
29	332.2	8.2	110000	2	EX004762_0 Mus muscu
30	332.2	8.2	110000	2	EX004762_1 Continuation (2 of
31	332.2	8.2	110261	10	AL807401 Mouse DNA
32	332.2	8.2	335799	2	EX005139 Mus muscu
33	297.4	7.3	315	4	AF301008 Bos tauru
34	289.4	7.1	110000	2	AL954766_2 Continuation (3 of
35	289.4	7.1	110000	2	EX004762_3 Continuation (4 of
36	253.8	6.2	982	6	AX364943 Sequence
37	250.8	6.2	2018	9	AF218896 Homo sapi
38	250.8	6.2	104907	9	AL132773 Human DNA
39	250.8	6.2	155869	2	AC023420 Homo sapi
40	250.8	6.2	167357	2	AC055771 Homo sapi
41	250.8	6.2	206647	9	AF002898 Homo sapi
42	246	6.0	246	6	AR273223 Sequence
43	246	6.0	246	6	AR276804 Sequence
44	246	6.0	246	6	AR407079 Sequence
45	246	6.0	246	6	AX368256 Sequence

ALIGNMENTS

RESULT 1  
AR164816  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL

AR164816  
Sequence 14 from patent US 6274339.  
AR164816  
AR164816.1 GI:16238025  
Unknown.  
Unclassified.  
1 (bases 1 to 8589)  
Moore, K. and Nagle, D. Lynn.  
Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity  
Patent: US 6274339-A 14 14-AUG-2001;  
PAT 17-OCT-2001

FEATURES	Location/Qualifiers
source	1. .8589
	/organism="unknown"
	/mol_type="unassigned DNA"
ORIGIN	
Query Match	98.0%; Score 3988.4; DB 6; Length 8589;
Best Local Similarity	99.0%; Pred. No. 0;
Matches 4013; Conservative 0; Mismatches 41; Indels 0; Gaps 0;	
QY	15 GCGCGCAACTGAGGCAAAAGCTTGAGGAGGAGGACGGCGCGACGCGACGCTCGCGGGCAG 74
DB	234 GGAGGCCGAGGCCGCTCGCGTGGCGCGCGCGGTGTCGGCTCGGCCGAGCCGAGGCCAA 293
QY	75 GAGCGGCGGGCCGCACTGTGTCAACGGCGGTGCGTGCAAACCTTGGCAACCGGCGAGTCGT 134
DB	294 GGAATGTGACCGGCCGTGTGTCAACGGCGGTGCGTGCAAACCTTGGCAACCGGCGAGTCGT 353
QY	135 CTGCCCCCGCGCTGGGTGGCGGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAAC 194
DB	354 CTGCCCCCGCGCTGGGTGGCGGAGCAATGCCAGCACTGGGGGGCGCTTCAGACTAAC 413
QY	195 TGGATCTTCTGGGTTTGTGACAGATGGACCTTGGAAATTATAAATACAAAACGAAGTCCAC 254
DB	414 TGGATCTTCTGGGTTTGTGACAGATGGACCTTGGAAATTATAAATACAAAACGAAGTCCAC 473
QY	255 GTGGCTCAITGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATCATTTTGTCTAC 314
DB	474 GTGGCTCAITGAAGGACAGCCAAATAGAAATATGAGACTTCGTTTCAATCATTTTGTCTAC 533
QY	315 AGAGTGTAGTTGGACCAATTATATGTGTTATGATGGGCACTCAATTTATGCACCGCTAGT 374
DB	534 AGAGTGTAGTTGGACCAATTATATGTGTTATGATGGGCACTCAATTTATGCACCGCTAGT 593
QY	375 TGCTGCATTTAGTGGCTCATTGTTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
DB	594 TGCTGCATTTAGTGGCTCATTGTTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 653
QY	435 TGTGTCCACATCAGGTTATGCTGTCTGCAATTTTTTTTATGATGATGCTGCTTATATTTGAC 494
DB	654 TGTGTCCACATCAGGTTATGCTGTCTGCAATTTTTTTTATGATGATGCTGCTTATATTTGAC 713
QY	495 TGGATTTTAATTAATCTTACAGTTTGTATGTGTCCAAATAACTGTCTCAGGCGAGAGAGA 554
DB	714 TGGATTTTAATTAATCTTACAGTTTGTATGTGTCCAAATAACTGTCTCAGGCGAGAGAGA 773
QY	555 GTGTAAAGATCAGTAATAGCAGCATCTGTTTGAATGTGAATGTTCTGAAACTGGAAAGG 614
DB	774 GTGTAAAGATCAGTAATAGCAGCATCTGTTTGAATGTGAATGTTTCTGAAACTGGAAAGG 833
QY	615 TGAAGCATGTGACATTCTCTACTGTACAGACAACTGTGGTTTTCTCATCGAGGCATCTG 674
DB	834 TGAAGCATGTGACATTCTCTACTGTACAGACAACTGTGGTTTTCTCATCGAGGCATCTG 893
QY	675 CAATTCAAGTGATGTGAGAGATGCTCTGCTTCTGAGATGGGATGCTTCTGAGATGTTTC 734
DB	894 CAATTCAAGTGATGTGAGAGATGCTCTGCTTCTGAGATGGGATGCTTCTGAGATGTTTC 953
QY	735 AGTTTCCTGTACCACTAACCACTCATTTTTCGAGCTCGAGAGGAATATTTCTTAACATTTAAAGCT 794
DB	954 AGTTTCCTGTACCACTAACCACTCATTTTTCGAGCTCGAGAGGAATATTTCTTAACATTTAAAGCT 1013
QY	795 CCCCAGAGCATCTCATAAAGCTGTGGTCAATTGGAAACATTATGTGGGTTGTGGAGGATA 854
DB	1014 CCCCAGAGCATCTCATAAAGCTGTGGTCAATTGGAAACATTATGTGGGTTGTGGAGGATA 1073
QY	855 TATGTTCAACCACTCAGATTATACATGGTTCTAGCGTATGACCTTGCCTTCTAGGGAGTG 914
DB	1074 TATGTTCAACCACTCAGATTATACATGGTTCTAGCGTATGACCTTGCCTTCTAGGGAGTG 1133
QY	915 GCTTCCACTAAACCGTTTCTGTGAAACAATGTGGTTTGTAGATATGGTCAATTTCTTTGGCAAT 974
DB	1134 GCTTCCACTAAACCGTTTCTGTGAAACAATGTGGTTTGTAGATATGGTCAATTTCTTTGGCAAT 1193

Qy	2055	GACAGCTGCAGGAGCTGTGCCCTGGACCAAGAACTGCCAGTGGAGAGCCCGGAATCAGGA	2114
Db	2274	GACAGCTGCAGGAGCTGTGCCCTGGACCAAGAACTGCCAGTGGAGAGCCCGGAATCAGGA	2333
Qy	2115	GTGCATTGCCCTGCCCGAAAATATCTCTGCAATTGGCTGGCATTTGGTTGGAAACTCATG	2174
Db	2334	GTGCATTGCCCTGCCCGAAAATATCTCTGCGAATTGGCTGGCAATTTGGTTGGAAACTCATG	2393
Qy	2175	TTTGAAAATTACTACTGCCAAGGAGAAATTATGACAAATGCTTAAATGTTCTGTAGGAACCA	2234
Db	2394	TTTGAAAATTACTACTGCCAAGGAGAAATTATGACAAATGCTTAAATGTTCTGTAGGAACCA	2453
Qy	2235	CAATGCCCTTTTGGCTTCTCTTACAACCCAGAGAAGGTAGAAATTTGCTCCTTAAGCAGCT	2294
Db	2454	CAATGCCCTTTTGGCTTCTCTTACAACCCAGAGAAGGTAGAAATTTGCTCCTTAAGCAGCT	2513
Qy	2295	GCGAATAATCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCATGGGTGGGCCCT	2354
Db	2514	GCGAATAATCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCATGGGTGGGCCCT	2573
Qy	2355	TCGGAAGATCAATGTGTCTTACTTGTTGCTGGGAAGATATGTCCCCATTTTACAATAGTTT	2414
Db	2574	TCGGAAGATCAATGTGTCTTACTTGTTGCTGGGAAGATATGTCCCCATTTTACAATAGTTT	2633
Qy	2415	ACTACAGTGAATGCCGTCTGAGCCAGTGTCTGTGGATTCTGTGGAAATTTTATCAGAACC	2474
Db	2634	ACTACAGTGAATGCCGTCTGAGCCAGTGTCTGTGGATTCTGTGGAAATTTTATCAGAACC	2693
Qy	2475	CAGTACTCGGGGACTGAAGGGTGCAACTGTGATCAACCCACTCAATGGTAGTGTCTGTGA	2534
Db	2694	CAGTACTCGGGGACTGAAGGGTGCAACTGTGATCAACCCACTCAATGGTAGTGTCTGTGA	2753
Qy	2535	AAGSCCTGTCAAAACACAGTGTAAAGCAGTGCAGGACACCAATGTGCCCTTGAGGACAGCATG	2594
Db	2754	AAGSCCTGTCAAAACACAGTGTAAAGCAGTGCAGGACACCAATGTGCCCTTGAGGACAGCATG	2813
Qy	2595	TGGAGATTGCACACAGCGCAGCTCTGAGTGCATGTGTGCAGCAACAATGAAGCAGTGTGT	2654
Db	2814	TGGAGATTGCACACAGCGCAGCTCTGAGTGCATGTGTGCAGCAACAATGAAGCAGTGTGT	2873
Qy	2655	GGACTCAATGCCATATGTGGCCCTCCTTCCCTTTTGGCCAGTGTATGAAATGGTATACGAT	2714
Db	2874	GGACTCAATGCCATATGTGGCCCTCCTTCCCTTTTGGCCAGTGTATGAAATGGTATACGAT	2933
Qy	2715	GAGCACCTGCCCCCTCAAAATTTGTCAGGCTACTGTACCTGTAGTCAITTCGTTGGAGCA	2774
Db	2934	GAGCACCTGCCCCCTCAAAATTTGTCAGGCTACTGTACCTGTAGTCAITTCGTTGGAGCA	2993
Qy	2775	ACGAGCTGTGGCTGGTGTACTGATCCAGCAATACTTGGCNAAGGGAATATGCATAGAGGG	2834
Db	2994	ACGAGCTGTGGCTGGTGTACTGATCCAGCAATACTTGGCNAAGGGAATATGCATAGAGGG	3053
Qy	2835	TTCTCTATAAGGACACAGTGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCACA	2894
Db	3054	TTCTCTATAAGGACACAGTGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCACA	3113
Qy	2895	GCCCTCTGCTCAATTCAGACATGTGTCTAGAGGACAGCAGATACAATCTGGTCTTTTCATCA	2954
Db	3114	GCCCTCTGCTCAATTCAGACATGTGTCTAGAGGACAGCAGATACAATCTGGTCTTTTCATCA	3173
Qy	2955	CTGTCCAGCTTGGCAAATGCAACGCCCAACAGTAAATGTCATCAATCAGAGCATCTGTGAGAA	3014
Db	3174	CTGTCCAGCTTGGCAAATGCAACGCCCAACAGTAAATGTCATCAATCAGAGCATCTGTGAGAA	3233
Qy	3015	GTGTGAGAACTGCACACAGCAAGCACTGCGAAGCTGCAATATCTCGCTTCTACGGTGA	3074
Db	3234	GTGTGAGAACTGCACACAGCAAGCACTGCGAAGCTGCAATATCTCGCTTCTACGGTGA	3293
Qy	3075	TCCACCAATGGAGGGAATGTCAAGCCATGCAATGCAATGGGCACACGCTCTCTGTGCAA	3134
Db	3294	TCCACCAATGGAGGGAATGTCAAGCCATGCAATGCAATGGGCACACGCTCTCTGTGCAA	3353
Qy	3135	CACCAACACGGGCAAGTGCTTCTGCACCAACCAAGGGGCTCAAGGGGAGCAGTGCCAGCT	3194

[illegible]

RESULT, 2  
AF531101  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM



Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidea; Bovidae; Bovinae; Bos.

**REFERENCE**

1. COVATTA, G. W., AND DOUGLAS, D. S.  
1 (bases 1 to 4493)

REFERENCE  
1. (Pages 1-20-4-25)  
AUTHORS  
Graphodatskaya, D., Joerg, H. and Stranzinger, G.

AUTHORS	TITLE
Grapovatskaya, D.; Coerg, H. and Strauszinger, S.	Direct Submission

FREE  
 JOURNAL  
 Submitted (18-JUL-2002) Animal Sciences. ETH. Tannenstrasse 1.  
 DIRECT SUBMISSION

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2016. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

**FEATURES** Location/Qualifiers  
Zurich, CH 8032, Switzerland

FEATURES	LOCATION/
SOURCE	1 4493

SOURCE  
I. 4493  
/organism-"Bos taurus"

"END"=exit [w/

```
/mot_type="mrna"  
/ab_yref="F55001.9913"
```

```
/ad_xref="taxon:"  
/ch_xref="13"
```

CPD  
1288  
/chromosome

CDS  
52.4299  
/not a membrane form. membrane protein

```
/note="transmembrane
/codes starting
```

```

/codon_start=1
/codon_end=1
/codon_start=1

```

```
/product="attractin"
/product?id="B3W05445"
/product?id="B3W05445" 1"
```

```
/protein_id="AAM95445.1"
/db_xref="GI:233033200"
/db_xref="GI:233033200"
```

```

/translation="WVAATEATRELRVTVPATSRNGRRCDWDATRAGRPGICRGT
RUPRWRRLLLPILLPPWAAEAAAASVGSAAAEKCEKTCWLLIEGPNRIMELR
GOCVCPAGVGEQCCHGGRFRITGSSGRTIDGPGNYKYKTKTWLIEGPNRIMELR
FNHPATECWDHLYVVDGDSIYAPLYVAASGLIYIPERDQNETPEVVAATSGYALLPFF
SDAAYLTGNITINFDMPNCCSGRGECKISNSNTVQCESNKGWBEACDIPHCVN
NACPFHRTCNSSDPGRCCSCFSWQPGCSVPVPAPOFWTRREYSLDLKLPRASHAV
VNGNIMWVGYNFNSDYMVILAYDLASREMLALNRNSVVRVYHSGLSIALYODKRY
MYGGKIDSTGNITNELRVHINESWLLSPKAKEQYAVVGHSAHVTILKNGRVMVLI
IPGHCPLYGVISSVOEYDLDKNTWSTILHTQGLAQVGGYGHSSVYDRTKALYTHGGYK
ASGANKRADDLYRDYDRTQMTWIIKDSRFYRLTAVIYVSGTMYVFGGNSHNDISM
SHGAKFSDDFWAYIDICORWSVLPRLPDHHDVNRFGHSAVLYNRTMYVFGGNSLULL
SDITLFTSQCEAESEACLAAGPVGRCMWDTGSSQCVSWELAPAEKIKCEBFSK
RIPDHORCDQLTDCYCATNTANGCWNCHVCPTNHSCTEGQISIFKYDHCPCNDPMY
YCNKTKSCRSALDQNCQEPNQEICIALPENICGIGWHLVGNCSKLIITARTYDNA
KILSCRHNAPLASLTWTKQVVEFYKQILRINQSSQSMKSKTLLTPWVGLRKINYISYWCW
DMSPTNSSLQMPSPSDAGFCGIIIEPSTRLGKAATCINPLNGSVCRPNANHSKQ
CRTPCALRTACGCTSGSGSNMCKQVDNAYVAFPPFGCCMEWTMSQSCPPEN
CSGYCTCSHCEGQGCWCTDPSNTGKCIKIGSVKGPVPMQSGGTGNSYPOPLLNS
SMCLDSRSHYFTHCPACQNGSHSICINQSIKCEKCNLTTHGKHCETISGFGYDPTTL
GKXQCPRCNGHSLCNTWGTGFCFTTKGVKGCOLCEVENRYQCNPLKGTCTVYLLT
IDYQTFEISLQEDDRYYTALINFAVTPDEQNRDLDMFINASKNPNLTINWAASFSAGTQ
AGEMVPSVKNTIKEXYDSDFSNEKDFRHNPIITFYVSNFTWPFIQIAFSGNSFN
MDIAPQVFTVFPFSGLLIYVAAYVWKIKQSCWAASRRBEOILREMOQAPRFAASVNA
LTDEEPPDLIGSITIKPKPIALECPNGKPIALFPCNGKPIALFVRLPGLGLGIPPPGOSGLF
ASAIVDPI SOOMPYYKVEKSGAVENKQOPPAOPGTCTI"

```

## ORIGIN

Query Match	87.7%	Score	3566;	DB	4;	Length	4493;		
Best Local Similarity	92.6%	Pred. No.	0;						
Matches	3756;	Conservative	0;	Mismatches	295;	Indels	3;	Gaps	1;

  

QY	15	GGCGGCAACTGAGCGAAGGCTGAGGAGGAGCGCGGCGACGGCAGCGCTGCCGGGCGAG	74
Db	249	GGCGGCTGAGCGCGCGGCGGCGACGGCGCTCGGTGTGGGCTCAGCGCAGCCGAGGCCAA	308
QY	75	GAGCGGCGGCGCGCACTGTGTCAACGGCGGTCTCGCTGCACCCCTGGCAACCGGCCAGTCCGT	134
Db	309	GGAATGTGACCGGCGCCCTGTGTCAACGGCGCGCCGCTCAACCCCTGGCAACCGGCCAGTGTGT	368
QY	135	CTGCCCGCGCGCTGGTGGCGGAGCAAAATGCCAGCACTGCGGGGGCGCTTCAGACTTAAC	194
Db	369	CTGCCCGCGCGGCTGGTGGCGGAGCAAAATGCCAGCACTGCGGGGGCGCTTCAGACTTAAC	428
QY	195	TGGATCTCTCGGGTTTGTGACAGATGCACCTGGAAATTTATAATACAAACGAAGTCAC	254
Db	429	TGGATCTCTCGGATTTTATACAGATGCACCTGGAAATTTATAATATAAAACAAATGCAC	488
QY	255	GTGGCTCATTGAGGACAGCCAAATAGATAAATGAGACTTCGTTCAATCATTTTGGCTAC	314
Db	489	ATGGCTCATTTGAAGGACAGCCAAATAGATAAATGAGACTTCGTTTCAATCATTTTGGCTAC	548
QY	315	AGAGTGTAGTGGGACCATTTATATGTGTTTATGATGGGACTCAATTTATGCACCGGTAGT	374





QY 661 CATCGAGGCATCTGCAATTCAAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAG 720  
Db CATCGAGGCATCTGCAATTCAAGTGAATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAG 720  
QY 721 GGTCTCGGATGTTGAGTCTCTGTAACAGCTTAACAGCTCAATTTGGACTCGAGAGGAATAT 780  
Db GGTCTCGGATGTTGAGTCTCTGTAACAGCTTAACAGCTCAATTTGGACTCGAGAGGAATAT 780  
QY 781 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGCTCAATGGAACCAATTAATGTGG 840  
Db TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGCTCAATGGAACCAATTAATGTGG 840  
QY 841 GTTGTGGAGGATATATGTTTCAACCACTCAGATTAATAAGCTGTGCTTCTAGCCTATGACCTT 900  
Db GTTGTGGAGGATATATGTTTCAACCACTCAGATTAATAAGCTGTGCTTCTAGCCTATGACCTT 900  
QY 901 GCTTCTAGGAGTGGCTTCCACTTAAACCGTTCGTGAACAAATGTGCTGTGATAGATGCT 960  
Db GCTTCTAGGAGTGGCTTCCACTTAAACCGTTCGTGAACAAATGTGCTGTGATAGATGCT 960  
QY 961 CATCTTTGGCATATATCAAGGATAAAATTTAATGATGAGGAAAAATTTGATTCACCT 1020  
Db CATCTTTGGCATATATCAAGGATAAAATTTAATGATGAGGAAAAATTTGATTCACCT 1020  
QY 1021 GGGAAATGTAACCAATGAGTTGAGATTTTTCATAATGAGTCATGGGTGTTGTTG 1080  
Db GGGAAATGTAACCAATGAGTTGAGATTTTTCATAATGAGTCATGGGTGTTGTTGTTG 1080  
QY 1081 ACCCTTAAGCAAGAGGAGTATGAGTGGTGGGCACTCTGCAACATGTTTACACTG 1140  
Db ACCCTTAAGCAAGAGGAGTATGAGTGGTGGGCACTCTGCAACATGTTTACACTG 1140  
QY 1141 AAGAAATGCGAGTGGTTCATGCTGGTCACTTTTGGTCACTGCTCTCTATGGATATA 1200  
Db AAGAAATGCGAGTGGTTCATGCTGGTCACTTTTGGTCACTGCTCTCTATGGATATA 1200  
QY 1201 AGCAATGTCAGGAATATGATTTGGATAAGAAACACATGAGATATATACACACCCAGGCT 1260  
Db AGCAATGTCAGGAATATGATTTGGATAAGAAACACATGAGATATATACACACCCAGGCT 1260  
QY 1261 GCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTGTTTACACATAGGACCGAGGCCCTA 1320  
Db GCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTGTTTACACATAGGACCGAGGCCCTA 1320  
QY 1321 TAGCTTCATGCTGCTACAGGCTTTTACAGTCCCAATAAGTACCGGCTTGCAGATGATCTC 1380  
Db TAGCTTCATGCTGCTACAGGCTTTTACAGTCCCAATAAGTACCGGCTTGCAGATGATCTC 1380  
QY 1381 TACCGATATGATGTGGATACCCAGATGTGACCAATTTTAAAGGACAGCCGATTTTCCGT 1440  
Db TACCGATATGATGTGGATACCCAGATGTGACCAATTTTAAAGGACAGCCGATTTTCCGT 1440  
QY 1441 TACTGACACAGCTGTGATGAGTGGGAAACCATGCTGGTGTGTTGGGGAAACACAC 1500  
Db TACTGACACAGCTGTGATGAGTGGGAAACCATGCTGGTGTGTTGGGGAAACACAC 1500  
QY 1501 AATGACACATCTATGAGCCATGCGCCAAATGCTTCTTTCAGATTTTCAATGAGCTATGAC 1560  
Db AATGACACATCTATGAGCCATGCGCCAAATGCTTCTTTCAGATTTTCAATGAGCTATGAC 1560  
QY 1561 ATTGCTGTGACCGCTGCTGATGCTTCCAGACCTGATCTCCACATGATGATCAACAGA 1620  
Db ATTGCTGTGACCGCTGCTGATGCTTCCAGACCTGATCTCCACATGATGATCAACAGA 1620  
QY 1621 TTTGGCCATT - CAGCAGTCTTACACACACACCATGATGCTGGTGTGTTCAATAG 1679  
Db TTTGGCCATTCCAGCAGTCTTACACACACACCATGATGCTGGTGTGTTCAATAG 1679  
QY 1680 TCTCTCTCCTCAGCGACATCTGTTATTCACCTCGGAAACAGTGTGATCGGATCGAGTGA 1739  
Db TCTCTCTCCTCAGCGACATCTGTTATTCACCTCGGAAACAGTGTGATCGGATCGAGTGA 1739

QY 1740 AGCCGCTTGTGTTAGCAGGAGACCTGTTATTCGGTGTGTGTGGAACACAGGGTCTCTCA 1799  
Db AGCCGCTTGTGTTAGCAGGAGACCTGTTATTCGGTGTGTGTGGAACACAGGGTCTCTCA 1799  
QY 1800 GTGTATCTGCTGGCGCTGGCAACTGATGAACAAGAAAAGTTAAATCAGAAATGTTT 1859  
Db GTGTATCTGCTGGCGCTGGCAACTGATGAACAAGAAAAGTTAAATCAGAAATGTTT 1859  
QY 1860 TTCCAAAAGAACTCTTGACCATGACATGCAATGTCCTCCAGGATACCAAGTGTACAGCTG 1919  
Db TTCCAAAAGAACTCTTGACCATGACATGCAATGTCCTCCAGGATACCAAGTGTACAGCTG 1919  
QY 1920 AGCCAAACCAATGACTGCTGCAATGCAATGTCCTCCAGGATACCAAGTGTACAGCTG 1979  
Db AGCCAAACCAATGACTGCTGCAATGCAATGTCCTCCAGGATACCAAGTGTACAGCTG 1979  
QY 1980 CTCAAGAGGCGAGATCTCCATTTTGTAGATGAGAAATTTGCCCAAGGATAACCTATGTA 2039  
Db CTCAAGAGGCGAGATCTCCATTTTGTAGATGAGAAATTTGCCCAAGGATAACCTATGTA 2039  
QY 2040 CTACTGTAACAAGAGACAGCTGAGAGCTGTGCTGCAACCAAGTGTGCAAGTGTGGA 2099  
Db CTACTGTAACAAGAGACAGCTGAGAGCTGTGCTGCAACCAAGTGTGCAAGTGTGGA 2099  
QY 2100 GCGCCGGAATCAGAGTGCATTTGCCCGAAAAATATCTGTGCAATTTGCTGCAATTT 2159  
Db GCGCCGGAATCAGAGTGCATTTGCCCGAAAAATATCTGTGCAATTTGCTGCAATTT 2159  
QY 2160 GGTGGAACCTCATGTTTGAATAATTAATCTGCAAGAGAAATTTATGCAATGCTAAAT 2219  
Db GGTGGAACCTCATGTTTGAATAATTAATCTGCAAGAGAAATTTATGCAATGCTAAAT 2219  
QY 2220 GTTCTGTAGGAACACAAATGCTTTTGGCTTCTTACAAACCCAGAAAGGTAGAAAT 2279  
Db GTTCTGTAGGAACACAAATGCTTTTGGCTTCTTACAAACCCAGAAAGGTAGAAAT 2279  
QY 2280 TGTCTTAAAGCAGCTCGAATAATTCAGTCACTCTCAGAGCATGTCGAAGCTCACTTAAC 2339  
Db TGTCTTAAAGCAGCTCGAATAATTCAGTCACTCTCAGAGCATGTCGAAGCTCACTTAAC 2339  
QY 2340 CCCATGGCTGGCTTGGAGATCAATGCTGCTCTGCTGCTGGAGATATGTCCTCC 2399  
Db CCCATGGCTGGCTTGGAGATCAATGCTGCTCTGCTGCTGGAGATATGTCCTCC 2399  
QY 2400 ATTTACAAATAGTTTACTACAGTGCATGCTGAGCCAGTGCATGCTGATTTCTGTGG 2459  
Db ATTTACAAATAGTTTACTACAGTGCATGCTGAGCCAGTGCATGCTGATTTCTGTGG 2459  
QY 2460 AATTTTATCAGAACCCAGTACTCGGGGACTGAGGCTGCAACCTGCTGATCAACCCACTCA 2519  
Db AATTTTATCAGAACCCAGTACTCGGGGACTGAGGCTGCAACCTGCTGATCAACCCACTCA 2519  
QY 2520 TGGTATGCTCTGTGAAAGGCTGCAACCAAGTGTCTAAGCAGTCCCGACACCATGTGC 2579  
Db TGGTATGCTCTGTGAAAGGCTGCAACCAAGTGTCTAAGCAGTCCCGACACCATGTGC 2579  
QY 2580 CTTGAGGACAGCATGTGGAGATTCACAGCGGAGCTCTGAGTGCATGCTGGTGCAGCAA 2639  
Db CTTGAGGACAGCATGTGGAGATTCACAGCGGAGCTCTGAGTGCATGCTGGTGCAGCAA 2639  
QY 2640 CATGAAGCAGTGTGAGTCTCAATGCTATGCTGCTCTTCTTCTTGGCCAGTGTAT 2699  
Db CATGAAGCAGTGTGAGTCTCAATGCTATGCTGCTCTTCTTCTTGGCCAGTGTAT 2699  
QY 2700 GGAATGGTATACGATGAGCAGCTGCCCTCTGAAAAATTTGTTTCTACTGTACTGTAG 2759  
Db GGAATGGTATACGATGAGCAGCTGCCCTCTGAAAAATTTGTTTCTACTGTACTGTAG 2759  
QY 2760 TCAATGCTTGGAGCAACAGCTGTGGTGTGCTGCTGATCCAGCANTACTGCAAGG 2819  
Db TCAATGCTTGGAGCAACAGCTGTGGTGTGCTGCTGATCCAGCAATACTGCAAGG 2819  
QY 2820 GAAATGCATAGAGGTTTCTTATTAAGGACAGTGAAGTGTCTTCCGCAAGCCCTACAGG 2879











QY 888 AGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACGGTCTCTGTGAACAATGTGGT 947  
Db 1110 AGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGTCTCTGTGAACAATGTGGT 1169  
QY 948 TGTGTAGATATGGTCATCTCTTTGGCAATATACAAAGGATAAATTTTACATGTATGGAGGAAA 1007  
Db 1170 TGTGTAGATATGGTCATCTCTTTGGCAATATACAAAGGATAAATTTTACATGTATGGAGGAAA 1229  
QY 1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCATCATTAATGAGTC 1067  
Db 1230 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTCATCATTAATGAGTC 1289  
QY 1068 ATGGGTCTCTGTGACCCCTTAAGCCAAAGAGGAGTATGCAAGTGGTGGGCACTCTCCACA 1127  
Db 1290 ATGGGTCTCTGTGACCCCTTAAGCCAAAGAGGAGTATGCAAGTGGTGGGCACTCTCCACA 1349  
QY 1128 CATTTGTACTGAAGAAATGCCGAGTGGTCATCTGCTGGTCACTTTTGGTCACTGCCCTCT 1187  
Db 1350 CATTTGTACTGAAGAAATGCCGAGTGGTCATCTGCTGGTCACTTTTGGTCACTGCCCTCT 1409  
QY 1188 CTATGGATATATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247  
Db 1410 CTATGGATATATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469  
QY 1248 ACACACCCAGGTGCCCTTGTGCAAGGGGTTTACGGCCATAGCAGTGTTTACGACCATAG 1307  
Db 1470 ACACACCCAGGTGCCCTTGTGCAAGGGGTTTACGGCCATAGCAGTGTTTACGACCATAG 1529  
QY 1308 GACCAGGGCCCTATACGTTTCATGGTGGCTACAGGCTTTCAGTGCCAAATGATACCGGT 1367  
Db 1530 GACCAGGGCCCTATACGTTTCATGGTGGCTACAGGCTTTCAGTGCCAAATGATACCGGT 1589  
QY 1368 TGCAGATGATCTCTACCGATGATGTGGATACCCAGATGTGACCATCTTTAAGGACAG 1427  
Db 1590 TGCAGATGATCTCTACCGATGATGTGGATACCCAGATGTGACCATCTTTAAGGACAG 1649  
QY 1428 CCGATTTTCCGTTACTTGACACAGCTGTGATGATGAGTGAACCAATCGTGGTGTGG 1487  
Db 1650 CCGATTTTCCGTTACTTGACACAGCTGTGATGATGAGTGAACCAATCGTGGTGTGG 1709  
QY 1488 GGGAAACACACAAATGACACATCTATGAGCCATGAGCCCAATGCTTCTCTTCAGATTT 1547  
Db 1710 GGGAAACACACAAATGACACATCTATGAGCCATGAGCCCAATGCTTCTCTTCAGATTT 1769  
QY 1548 CATGGCCTATGACATTCCTGTGACCGCTGGTCAGTGTCTCCAGACCTGATCTCCACCA 1607  
Db 1770 CATGGCCTATGACATTCCTGTGACCGCTGGTCAGTGTCTCCAGACCTGATCTCCACCA 1829  
QY 1608 TGAATGCAACAGATTTGGCCATTCAGCAGTCTTTACACACAGCACCATGATGTGTCCG 1667  
Db 1830 TGAATGCAACAGATTTGGCCATTCAGCAGTCTTTACACACAGCACCATGATGTGTCCG 1889  
QY 1668 TGGTTTCAATAGTCTCTCTCAGCAGACATCTCTGGTATTCACCTCGGAACAGTGTGATGC 1727  
Db 1890 TGGTTTCAATAGTCTCTCTCAGCAGACATCTCTGGTATTCACCTCGGAACAGTGTGATGC 1949  
QY 1728 GCATCGAGTGAAGCCGCTGTTTACGAGCAGACCTGGTATTCGGTGTGTGTGGAACAC 1787  
Db 1950 GCATCGAGTGAAGCCGCTGTTTACGAGCAGACCTGGTATTCGGTGTGTGTGGAACAC 2009  
QY 1788 AGGTCTCTCAGTGTATCTCTGGGGCTGGCAACTGATGAACAAGAAAGTTAAA 1847  
Db 2010 AGGTCTCTCAGTGTATCTCTGGGGCTGGCAACTGATGAACAAGAAAGTTAAA 2069  
QY 1848 ATCAGAAATGTTTTCCAAAAGAACTCTTGACCAATGACAGATGTGACGACACAGATTG 1907  
Db 2070 ATCAGAAATGTTTTCCAAAAGAACTCTTGACCAATGACAGATGTGACGACACAGATTG 2129  
QY 1908 TTACAGCTGACAGCCCAACCAATGATGCTGCCACTGGTGCATGCAATGTGTCCCCAG 1967  
Db 2130 TTACAGCTGACAGCCCAACCAATGATGCTGCCACTGGTGCATGCAATGTGTCCCCAG 2189

QY 1968 GNAACACAGCTGCTCAGAGGCCAGATCTCCATTTTATAGTATGAGAAATTTGCCCCNAGGA 2027  
Db 2190 GNAACACAGCTGCTCAGAGGCCAGATCTCCATTTTATAGTATGAGAAATTTGCCCCNAGGA 2249  
QY 2028 TAAACCTATGTACTACTGTAAACAAGAACACAGCTGCAAGGAGCTGTGCCCTGGACAGAA 2087  
Db 2250 TAAACCTATGTACTACTGTAAACAAGAACACAGCTGCAAGGAGCTGTGCCCTGGACAGAA 2309  
QY 2088 CTGCAGGTGGAGCCGCCGAATCAGGAGTCATTTGCCCTGCCGGAATATCTGTGGCAT 2147  
Db 2310 CTGCAGGTGGAGCCGCCGAATCAGGAGTCATTTGCCCTGCCGGAATATCTGTGGCAT 2369  
QY 2148 TGGCTGGCAATTTGGTGTGAAAATCTATGTTGAAAATTAATACTGTCGAAGGAGAATTATGA 2207  
Db 2370 TGGCTGGCAATTTGGTGTGAAAATCTATGTTGAAAATTAATACTGTCGAAGGAGAATTATGA 2429  
QY 2208 CAATGCTAAATTTGTTCTGTAGGAACCAATAGCCCTTTTGGGCTTCTTTAACAACAGAA 2267  
Db 2430 CAATGCTAAATTTGTTCTGTAGGAACCAATAGCCCTTTTGGGCTTCTTTAACAACAGAA 2489  
QY 2268 GAAGGTAGAAATTTGTCTTAAGCAGCTGCGAATATGCAATCATCTCAGACATGTCCAA 2327  
Db 2490 GAAGGTAGAAATTTGTCTTAAGCAGCTGCGAATATGCAATCATCTCAGACATGTCCAA 2549  
QY 2328 GCTACCTTAACCCCATCGGTGCGCTTCGGAAGATCAATGTGCTCTACTGCTGTGGGA 2387  
Db 2550 GCTACCTTAACCCCATCGGTGCGCTTCGGAAGATCAATGTGCTCTACTGCTGTGGGA 2609  
QY 2388 AGATATGTCGCCANTTTACAAATAGTTTACTACAGTGGATGCGGCTGAGCCAGTGATGC 2447  
Db 2610 AGATATGTCGCCANTTTACAAATAGTTTACTACAGTGGATGCGGCTGAGCCAGTGATGC 2669  
QY 2448 TGGATCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGGAAGGCTGCAACCTGCAT 2507  
Db 2670 TGGATCTGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGGAAGGCTGCAACCTGCAT 2729  
QY 2508 CAAACCCATCAATGTGTGTGTGAAAGGCTGCAACCAACACAGTGTCTAGCAGTGGCG 2567  
Db 2730 CAAACCCATCAATGTGTGTGTGAAAGGCTGCAACCAACACAGTGTCTAGCAGTGGCG 2789  
QY 2568 GACACCATGTGCTTGGAGACAGCATGTGGAGATTGCAACACGCGGACGCTCTGATGTCAT 2627  
Db 2790 GACACCATGTGCTTGGAGACAGCATGTGGAGATTGCAACACGCGGACGCTCTGATGTCAT 2849  
QY 2628 GTGGTGCAGCAACATGAAGCAGTGTGTGCACTCCAAATGCCCTATGTGGCTCTTCCCTTT 2687  
Db 2850 GTGGTGCAGCAACATGAAGCAGTGTGTGCACTCCAAATGCCCTATGTGGCTCTTCCCTTT 2909  
QY 2688 TGGCCAGTGTATGGAATGTTATACGATGAGCACCTGCCCTCCCTGAAATTTGTCAGGCTA 2747  
Db 2910 TGGCCAGTGTATGGAATGTTATACGATGAGCACCTGCCCTCCCTGAAATTTGTCAGGCTA 2969  
QY 2748 CTGTACCTGTAGTCAATTCCTTGGAGCAACCCAGGCTGTGGCTGTGTACTGTATCCAGCAA 2807  
Db 2970 CTGTACCTGTAGTCAATTCCTTGGAGCAACCCAGGCTGTGGCTGTGTACTGTATCCAGCAA 3029  
QY 2808 TACTGGCAAAAGGAAATGCATAGAGGTTCTTATAAGGACACAGTGAAGATGCCCTTCGCA 2867  
Db 3030 TACTGGCAAAAGGAAATGCATAGAGGTTCTTATAAGGACACAGTGAAGATGCCCTTCGCA 3089  
QY 2868 AGCCCTCTACAGAAATTTCTATCCACAGCCCTGCTCAATTCACAGCATGTGTCTAGAGGA 2927  
Db 3090 AGCCCTCTACAGAAATTTCTATCCACAGCCCTGCTCAATTCACAGCATGTGTCTAGAGGA 3149  
QY 2928 CAGCAGATACAACTGGTCTTTTCATTCATCTGTCAGCTTGGCAATGCAACGCCACAGTAA 2987  
Db 3150 CAGCAGATACAACTGGTCTTTTCATTCATCTGTCAGCTTGGCAATGCAACGCCACAGTAA 3209  
QY 2988 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACACCAAGCAAGCAAGTGGGA 3047  
Db 3210 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACTGACACCAAGCAAGCAAGTGGGA 3269  
QY 3048 GACCTGCATATCTGGCTTCTTACGGTGTATCCCAACCAATGGAGGGAATGTGACCCATGCAA 3107

Db	3270	GACCTGATATCTGGCTTCTACGGTGTATCCCAATGGAGGGAATGTGAGCATGCAC	3329
Qy	3108	GTGCAATGGGACACCGCTCTGTGTCAACACCAACACCGGCAAGTGTCTTGACACCA	3167
Db	3330	GTGCAATGGGACACCGCTCTGTGTCAACACCAACACCGGCAAGTGTCTTGACACCA	3389
Qy	3168	GGGCGTCAAGGGGACAGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGAAACCC	3227
Db	3390	GGGCGTCAAGGGGACAGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGAAACCC	3449
Qy	3228	TCTCAGAGGAACATGTTATATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATC	3287
Db	3450	TCTCAGAGGAACATGTTATATATCTCTCTTATTGACTATCAGTTCACCTTTAGTCTATC	3509
Qy	3288	CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTGACGAAACAA	3347
Db	3510	CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTGACGAAACAA	3569
Qy	3348	CAGGATTTGACATGTTTCATCATGCTCCACGAATTTTCAACCTCAACATCACTGGGC	3407
Db	3570	CAGGATTTGACATGTTTCATCATGCTCCACGAATTTTCAACCTCAACATCACTGGGC	3629
Qy	3408	TGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATCCCTGTTGTTTCAAAAACCA	3467
Db	3630	TGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATCCCTGTTGTTTCAAAAACCA	3689
Qy	3468	CATTAGAGGTACAAAGATGTTTCTCTATGAGAGTTTGATTTGCGAACCAACCCAA	3527
Db	3690	CATTAGAGGTACAAAGATGTTTCTCTATGAGAGTTTGATTTGCGAACCAACCCAA	3749
Qy	3528	TATCACTTCTCTTGTATGTGCTAATTTTCACTGCGCCCATCAAAATTCAGAT	3581
Db	3750	TATCACTTCTCTTGTATGTGCTAATTTTCACTGCGCCCATCAAAATTCAGAT	3803

RESULT 6  
AB062913  
LOCUS AB062913 Mesocricetus auratus Atrn mRNA linear ROD 16-JAN-2002  
DEFINITION Mesocricetus auratus Atrn mRNA for attractin, complete cds.  
ACCESSION AB062913  
VERSION AB062913.1 GI:16930100  
KEYWORDS  
SOURCE Mesocricetus auratus (golden hamster)  
ORGANISM Mesocricetus auratus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Mesocricetus.  
1 Kuramoto,T., Nomoto,T., Fujiwara,A., Mizutani,M., Sugimura,T. and Ushijima,T.  
Insertional mutation of the Attractin gene in the black tremor hamster  
Mamm. Genome 13 (1), 36-40 (2002)  
21635551  
11773967  
2 (bases 1 to 4284)  
Kuramoto,T. and Ushijima,T.  
Direct Submission  
Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramotogan2.ncc.go.jp, URL:http://www.ncc.go.jp/research/rat-genome/, Tel:81-3-3542-2511, Fax:81-3-5565-1753)  
Location/Qualifiers  
1. .4284  
/organism="Mesocricetus auratus"  
/mol type="mRNA"  
/db xref="taxon:10036"  
1. .4284  
/gene="Atrn"  
1. .4284  
/gene="Atrn"  
FEATURES  
source  
gene  
CDS

/codon_start=1	
/product="attractin"	
/protein_id="BAB72012.1"	
/translation="MWAAAAAALLRSTAAATAAPAGRKVEQVRLCAATGAWRPGPR	
AGLCLPVWSPALPPLLLLLPLPREAAASAAVSSGSAAYKTKCTWLI	
CVNGRCNPGTGCQCVAGWEOCHQCGFRLTGSSGFVTDGPGNYKTKCTWLI	
EGQPNRLWRLFNHFAFECSDHLYVYDGSIVAPLIAESGLIIPERDNETAPEVT	
VTSGLLHFFSDAAYNLGNITNFDMCPNCSGRGECSSNITVCESENWK	
GESDIPECTDNCGPHRGICNASDVSCSPHWGPQCSIPVPAQNSFWIREYED	
FKLPRASHKAVNGNIMWVGYNFNHSDYSLAYDLAREMLPLNHSVNVVVRVG	
HSLLAHKDKIMYGCKIDSTGNVTNELRVFHINESWLLTPKAKEQYAVGHSAHV	
TLATRVMLVIFGHCPYLYGISVQVQEDLEKNTWSILOQALVQGGYGVHSDHVR	
TKALVNGYKAFSANKYRLADLYRYVDVDTOMWTLILKDSFRFLYLTAVTSYMLV	
FGNTHDTSNHSAGCFSDPMAYDIACDPMVLPRLDLDHVDNRHGHSAVLHNSM	
YVGFNSLLSLDVLVFTSECDHAHSEACVAAGPGRICLMDTVSSRCTSWELATEE	
QAEKSECFKRLTDHRCQHTDYSCTANTINDCHCNDHCVFVNHSCTEGGISTS	
KYDNCPKDNPMYCNKTKSCRCALDQNCQWEPNOECIALPENICGHWLGNLSCL	
KITAKENYDNALSCRNHAFSLASOKKFEVLKHLRMSQSSMSKLTITPWVG	
LRKTNVSWCWEDMSPTNLSLLOWPSEPSPDAGFCGLISPTSLGKAACTINPLNGS	
VCEPANESAKOCTPCALRTACSECTSSSECMWCMQCDVSNAYVASFPQGM	
EWYTMSCPPENCSEYCTCSHCLRPQCGWCTDPNTGKGLGKLEGSYKPIRMPQAS	
TGNTVFPQLNNSMCLDSRYNWSFIHCPACQCNHCKSCINQSCICEKNLTITGKHOE	
TCISGFYDPTNGKQPCCKNGHASLNTINTGKCTCTTGKVGKDEQOLCEVENRYOG	
NPLKGTCTYTLIDYQTFSLQEDDRYTAINFVATPDQNRDLDMFINAKSNFNLN	
ITWATSFAGTQAGEMPVWSKTNKEYKDSFNSKEPFRNHNITFFVYSNFTWPI	
KIQAFSQHSNFMDLVQFVFTFSCFLSLLLVAAVWVKIKQSCWASRRQLRMOQ	
MASPPASVNVALSTDEPDLIGSITKVPKPIALEPCFNGKAAVLSVFLPRGLG	
GIPPGQSLAVASALVDISQMPFVYKESGAVNRKQPPPAQPGTCI"	
ORIGIN	
Query Match	83.5%; Score 3395.4; DB 10; Length 4284;
Best Local Similarity	89.9%; Pred. No. 0;
Matches 3642; Conservative	0; Mismatches 411; Indels 0; Gaps 0;
Qy	16 CGCGCACTGAGGCAAGCTCAGGAGGAGGAGCGCGGCGGCGGCGGCGGCGGCGG 75
Db	232 GAGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCAAG 291
Qy	76 AGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGTC 135
Db	292 GAATGTGACCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGTC 351
Qy	136 TGCCCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTACT 195
Db	352 TGCCCCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTACT 411
Qy	196 GGATCTCTTGGGTTGTGACAGATGGACCTCGGAAATTTATAAATACAAACGAACTGACG 255
Db	412 GGATCTCTTGGATTCGTACAGATGGACCTGGAAATTTATAAGTATAAACAAGTGCACA 471
Qy	256 TGGCTCATTTGAAGACAGCCAAATAGAAATAGAGCTCGTTTCAATCATTTTGTCTACA 315
Db	472 TGGCTCATTTGAAGACAGCCAAATAGAAATAGAGCTCGTTTCAATCATTTTGTCTACA 531
Qy	316 GAGTGTAGTGGGACCAATTTATATGTTTATGATGGGACTCAATTTTATGACCGCTAGTT 375
Db	532 GAATGTAGTGGGACCAATTTATATGTTTATGATGGGACTCAATTTTATGACCTCTGATT 591
Qy	376 GGTGCAATTTAGTGGCCTCATTTGTTCTTGAGAGAGATGCAATGAGCTGTCCCTGAGGTT 435
Db	592 GGTGCTTTTAGTGGTCTCATTTGTTCTTGAGAGAGATGCAATGAGAGCTGAGAGCTGAGAG 651
Qy	436 GTTGGCAGATCAGGTTATGCTGCTGCTATTTTATGATGATGCTGCTTATTAATTTGACT 495
Db	652 ACTGTCACTTCAAGTTATGCTGCTGCTATTTTCAAGTATGCTGCTTATTAATTTGACT 711
Qy	496 GGATTTAATATTTACTTACAGTTTGTGATAATGTGCTCAAAATAAATCTCTCAGGCGGAGAGAG 555
Db	712 GGATTTAATATTTACTTACAGTTTGTGATAATGTGCTCAAAATAAATCTCTCAGGCGGAGAGAG 771
Qy	556 TGTAGATCAGTAAATAGCAGGATCTGTTGAATGTGATGCTGCTGCTGCTGCTGCTGCTGCT 615

Db 772 TGTAAGAGCAGTAACAGCAGCAATACATGTTGAGTGTGAATGTTCTGAAAACTGGAAAGGG 831  
Qy 616 GAACAGATGTGACATTTCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCATCTGC 675  
Db 832 GAATCGTGTGACATTTCCCATTTGTACAGACAACTGTGGTTTTCTCATCCGAGGCATCTGT 891  
Qy 676 AATTCAAGTGAATGTACAGAGATGCTGCTGTTCTCAGACTGGCAGGTCCTGATGTTCA 735  
Db 892 AATGCAAGTGAACGTGAGAGGGTGTCTGCTCCCGCACTGGCAGGGTCTCGATGTTCA 951  
Qy 736 GTTCTCTGACCACTAACAGTCAATTTGGACTCGAGAGGAATATTCTAACTTAAAGCTC 795  
Db 952 AATTCCTGTGCCACTAACAGTCTTTCTGGAATTCGAGAAGATATTCCGATTTCAAGCTC 1011  
Qy 796 CCAGAGCATCTCATAAAGCTGTGGTCAATGGAACAAATATTGTGGTGTGTTGAGGAATAT 855  
Db 1012 CCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAATATAATGTGGTGTGTTGTTGGATAT 1071  
Qy 856 ATGTTCAACCACTCAGATATACATGTTTCTAGCGTATGACCTGCTCTTAGGAGTGG 915  
Db 1072 ATGTTCAACCACTCAGATATACATGTTTCTAGCGTATGACCTGCTCTTAGGAGTGG 1131  
Qy 916 CTTCCACTAAACCGTTCTGTGAACAAATGTGGTGTGATATGTTCTTTTGGCATTA 975  
Db 1132 CTTCCGCTAAACCACTCTGTGAACAAATGTGGTGTGATATGTTCTTTTGGCATTA 1191  
Qy 976 TACAAGGATAAATTTACATGATGAGGAAATTTGATTTCACTGGGAATGTGACCAAT 1035  
Db 1192 CATAGGATAAATTTACATGATGAGGAAATTTGATTTCAACAGGGAATGTGACCAAT 1251  
Qy 1036 GAGTTGAGAGTGTGTTTTCACATTTCAATATGAGTCAATGGTGTGTTGACCCCTAAGGCNAAG 1095  
Db 1252 GAGCTGAGAGTATTTTCATATTCATATGAATCATGGTATGTTTAAACACCAAGCAAG 1311  
Qy 1096 GAGCAGTATGCAAGTGTGGGACTCTGACACATTTGTTTACACTGGAATGGCGAGTG 1155  
Db 1312 GAGCAGTATGCAAGTGTGGGACTCTGACACATTTGTTTACACTGGAATGGCGAGTG 1371  
Qy 1156 GTCATGCTGTCATCTTTGTCATCTGCTCTCTATGGATATATAGCAATGTGACGAA 1215  
Db 1372 GTCATGTTGTCATCTTTGTCATCTGCTCTCTATGGATATATAGCAATGTGACGAA 1431  
Qy 1216 TATGATTTGGATTAAGAACATCGAGTATATTAACACCCAGGTCCTTTGTGCAAGG 1275  
Db 1432 TATGACTTGAAGAAAGAACATCGAGTATATTAACAACTCAGGGTCTCTTGTGCAAGGA 1491  
Qy 1276 GGTACGGCCATAGCAGTGTGTTACGACCATAGACCCAGGCCCCCTATACGTTTATGGTGGC 1335  
Db 1492 GGTATGGCCACAGCAGTGTGTTATGATCACAGACCAAGGCCCTGTATGTTTATGGTGGC 1551  
Qy 1336 TACAAGGCTTTCAAGTCCCAATAGTACCGGCTTTGCAAGTGTCTTACCGCATATGATGTG 1395  
Db 1552 TACAAGGCTTTTACGCCCAATAAATACCGGCTTTGCAAGTGTCTTACAGATATGATGTG 1611  
Qy 1396 GATACCGAGATGTGAGACCATTTTAAGGACAGCCGATTTTTCGGTTACTTGCACAGCT 1455  
Db 1612 GATACCGAGATGTGAGACCATTTTAAGGACAGCCGATTTTTCGGTTACTTGCATACAGCT 1671  
Qy 1456 GTCATAGTGTAGTGAACCACTGTGGTGTGTTGGGGAAACACACACATGACATCTATG 1515  
Db 1672 GTCATAGTGTAGTGAACCACTGTGGTGTGTTGGGGAAACACACACATGACATCTATG 1731  
Qy 1516 AGCCATGGGCCCAAAATGCTTCTCTTACAGATTTTATGGCCTATGACATTCCTGTGACCGC 1575  
Db 1732 AGCCATGGGCCCAAAATGCTTCTCTTACAGATTTTATGGCCTATGACATTCCTGTGACCGA 1791  
Qy 1576 TGTGTAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGCGCATTCAGCA 1635  
Db 1792 TGTGTAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGCGCATTCAGCA 1851  
Qy 1636 GTCTTACACAAACAGCACCATGTATGTTGTTGTTTCAATAGTCTCTCTCTCAGCGAC 1695  
Db 1852 GTCTTGCACAAACAGTACCAATGATGTTTTCGGCGGTTTCAACAGACCTCTCTCTCAGTGAC 1911

Qy 1696 ATCCTGTTATTTCACTCGGAAACAGTGTGATCGCATCGAGTGAAGCGCTTGTGTTAGCA 1755  
Db 1912 GTCTTTAGTCTTTTACTCGGAGCAATGTGACGCCACCACACAGTGGAGCGCTTGTGTAGCA 1971  
Qy 1756 GCAGGACCTGTGATTTCCGTTGTGTGGAAACACAGGTCCTCTCAGTGTATCTCTGTTGGCG 1815  
Db 1972 GCAGGACCTGTGATTTCCGTTGTGTGGGACACAGTGTCACTCTCGATGTACCTCTGTTGGAG 2031  
Qy 1816 CTGGCAACTGTATGAACCAAGAAAGAAAGTTAAATCAGAATGTTTTCCTCAAAAGAACTCTT 1875  
Db 2032 TTGGCACTGAAGAAACAGCAGAAAGTTAAATCAGAATGTTTTCCTCAAAAGAACTCTT 2091  
Qy 1876 GACCATGACAGATGTGACAGCAGACACAGATTTGTACAGTGCACAGCCAAACACCAATGAC 1935  
Db 2092 GACCATGACAGATGTGACAGCAGACACAGATTTGTACAGTGCACAGCCAAATACCAATGAC 2151  
Qy 1936 TGCCACTGGTGCATGACCAATGTGTCCTGAGGAAACACAGCTGCTCAGAAAGGCGCAGATC 1995  
Db 2152 TGCCACTGGTGCATGACCAATGTGTCCTGAGGAAACACAGCTGCAAGAGGCGCAGATC 2211  
Qy 1996 TCCATTTTGGTATGAGAATTTGCCCAAGGAATTCCTATGTACTACTGTAAACAAGAA 2055  
Db 2212 TCCATTTTCCAAGTATGACAAATTTGCCCAAGGAATTCCTATGTACTACTGTAAACAAGAA 2271  
Qy 2056 ACCAGCTGCAGAGCTGTGCTCGACAGCAACTGCCAGTGGGAGCCCGGAAATCAGAG 2115  
Db 2272 ACAGCTGCAGAGCTGTGCTCGACAGCAACTGCCAGTGGGAGCCCGGAAATCAGAG 2331  
Qy 2116 TGCAATGCTCTGCCGCAAAATATCTGTGGCAATTTGGCTGGCATTTTGGTAACTCATGT 2175  
Db 2332 TGCAATGCTCTGCCGCAAAATATCTGTGGCAATTTGGCTGGCATTTTGGTAACTCATGT 2391  
Qy 2176 TTGAAATTAATCTACTGCCAAGGAAATATGACAAATGCTAAATTTGTTCTGTAGAACCCAC 2235  
Db 2392 CTGAAATTAATCTACTGCCAAGGAAATATGACAAATGCTAAATTTGTTCTGTAGAACCCAC 2451  
Qy 2236 AATGCCCTTTTGGCTTCTCTTACAAACCAAGAAAGTAGAATTTGCTTTAAGCAGCTG 2295  
Db 2452 AATGCCCTTTTGGCTTCTCTTACAAACCAAGAAAGTAGAATTTGCTTTAAGCAGCTT 2511  
Qy 2296 CGAATAATGCAAGTCAATCTCAGAGCATGTCCAAAGTCACTTTAACCCCATGGTGGCGCTT 2355  
Db 2512 CGAATAATGCAAGTCAATCTCAGAGCATGTCCAAAGTCACTTTAAGCTTCCATGGTGGCTT 2571  
Qy 2356 CGGAAGATCAATGTGCTCTGCTGCTGGGAAGATATGCTCCCATTTTACAAATAGTTTA 2415  
Db 2572 CGGAAGATCAATGTGCTCTGCTGCTGGGAAGATATGCTCCCATTTTACAAATAGTTTG 2631  
Qy 2416 CTACAGTGGATGCCGTCTGAGCCCAAGTCAATGCTGGATTTCTGTGAAATTTTATCAGAACCC 2475  
Db 2632 CTGCAAGTGGATGCCCATCGAGCCCAAGTCAATGCTGGATTTCTGTGGCATCTTGTGAGGCT 2691  
Qy 2476 AGTACTGGGACCTGAAGGCTGCAACCTGCAATCAACCCACTCAATGTTGTGTCTGTGAA 2535  
Db 2692 AGTACTGGGACCTGAAGGCTGCAACCTGCAATCAACCCACTCAATGTTGTGTGTGAA 2751  
Qy 2536 AGGCTGTGCAAAACCAAGTCTTAAAGCAGTGGCCGACACCATGTGCTTTCAGACAGCATGT 2595  
Db 2752 AGGCTGTGCAAAACCAAGTCTTAAAGCAGTGGCCGACACCATGTGCTTTCAGACAGCATGT 2811  
Qy 2596 GGAGATGCAACAGCGGAGCTCTGAGTGCATGTTGGTGGAGCAACATGAAGCAGTGTGTG 2655  
Db 2812 AGTGAATGCAACAGCGGAGCTCTGAGTGCATGTTGGTGGAGCAACATGAAGCAGTGTGTG 2871  
Qy 2656 GACTCCAATGCTATGCGCTCTCTCCCTTTTGGCCAGTGTATGGAATGTATACCATG 2715  
Db 2872 GACTCCAATGCTATGCGCTCTCTCCCTTTTGGCCAGTGTATGGAATGTATACCATG 2931  
Qy 2716 AGCACTGTCGCCCTGAAAAATTTGTTCAAGGCTATGTTACCTGTATGTTCTGTGGAGCA 2775  
Db 2932 AGCACTGTCGCCCTGAAAAATTTGTTCTGGCTATGTTACCTGTATGTTCTGTGGAGCAG 2991





1995 CTCCATTTTATGATGAGAAATGCCCCAAGGATACCTTATCTACTTACTGTATCAAGAA 2054  
2283 CTCATTTGCCAAGATGAGAGTTGCCCCAAGGATACCTTATCTACTTACTGTATCAAGAA 2342  
2055 GACGAGCTGACGAGCTGTGCTTGGACAGAACTGACAGTGGGAGCCCGGAATCAGGA 2114  
2343 AACGAGCTGACGAGCTGTGCTTGGACAGAACTGACAGTGGGAGCCCGGAATCAGGA 2402  
2115 GTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2174  
2403 GTGCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2462  
2175 TTTGAAATTTACTACTGCCAAGAGAAATATGCAATGCTTAAATTTGCTGTAGGAACCA 2234  
2463 TCTGAAATTTACTACTGCCAAGAGAAATATGCAATGCTTAAATTTGCTGTAGGAACCA 2522  
2235 CAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2294  
2523 CAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2582  
2295 GCGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2354  
2583 TCGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2642  
2355 TCGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2414  
2643 TCGAATTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2702  
2415 ACTACAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2474  
2703 GCTGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2762  
2475 CAGTACTCGGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2534  
2763 TAGTACTCGGAGTGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2822  
2535 AAGGCTGCAAAACACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2594  
2823 AAGGCTGCAAAACACAGTGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2882  
2595 TGGAGATGCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2654  
2883 TGGAGATGCAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2942  
2655 GGAATCCAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2714  
2943 GGAATCCAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3002  
2715 GAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2774  
3003 GAGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3062  
2775 ACCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2834  
3063 GCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3122  
2835 TTCTATTAAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2894  
3123 CAGCTATAAAGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3182  
2895 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2954  
3183 GCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3242  
2955 CTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3014  
3243 CTGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3302  
3015 GTGTGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3074  
3303 GTGTGAGAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3362

3075 TCCACCAATGAGGAAATGTCAGCCATGCAATGCAATGGGCAACGCTCTCTGTGCAA 3134  
3363 CCCGACTANTGGAGCAATGTCAGCCATGCAATGGGCAACGCTCTCTGTGCAA 3422  
3135 CACCAACACGGGCAAGTGTCTGTGCAACCAAGGGGCTCAAGGGGACAGTGTGCAAGCT 3194  
3423 CACCAACACGGGCAAGTGTCTGTGCAACCAAGGGGCTCAAGGGGACAGTGTGCAAGCT 3482  
3195 ATGTGAGTACAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTTATCTCT 3254  
3483 ATGTGAGTACAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATTTATCTCT 3542  
3255 TCTTATTCAGCTATCAGTTTCACTTGTAGTCTATCCAGGAAAGATGCTGCTTATTCAGC 3314  
3543 TCTTATTCAGCTATCAGTTTCACTTGTAGTCTATCCAGGAAAGATGCTGCTTATTCAGC 3602  
3315 TATCAATTTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3374  
3603 CATCAACTTTTGTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3662  
3375 CTCCAAAGAAATTTCAACCTTCAACATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3434  
3663 CTCCAAAGAAATTTCAACCTTCAACATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3722  
3435 TGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3494  
3723 TGGAGAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3782  
3495 TAAATGAGAAATTTGATTTTTCGCAACCAACCTTCAACATCAGCTGCTGCTGCTGCTGCTGCTG 3554  
3783 TAAATGAGAAATTTGATTTTTCGCAACCAACCTTCAACATCAGCTGCTGCTGCTGCTGCTGCTG 3842  
3555 TTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3614  
3843 TTTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3902  
3615 GGTACAGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3674  
3903 GGTACAGTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3962  
3675 GGTGAGAGATCAACAAAGTCTTGGGCTTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3734  
3963 GGTGAGAGATCAACAAAGTCTTGGGCTTCCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4022  
3735 GATGCAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3794  
4023 GATGCAACAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4082  
3795 GGAGCTCTCTGATCTTATTTGGGGGAGTATAAAGACTGTTTCCCAAAACCTTTCGACTGGA 3854  
4083 AGAATCTCTGATCTTATTTGGGGGAGTATAAAGACTGTTTCCCAAAACCTTTCGACTGGA 4142  
3855 GCGGTGTTTGGGCAACAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3914  
4143 GCGGTGTTTGGGCAACAAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4202  
3915 GGGTGGCACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3973  
4203 GGGGAGAAATCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4262  
3974 TTTTCTCAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4033  
4263 TTTTCTCAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4322  
4034 AGCAGCCCTTGCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4068  
4323 AGCAGCCCTTGCACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4357

RESULT 8  
AF119821  
LOCUS  
DEFINITION Mus musculus attractin (Mgca) mRNA, complete cds.

AF119821 4313 bp mRNA linear ROD 16-APR-1999  
Mus musculus attractin (Mgca) mRNA, complete cds.



[illegible]

## ORIGIN

	Query Match	82.5%;	Score 3354.8;	DB 10;	Length 4313;				
	Best Local Similarity	89.2%;	Pred. No. 0;	Mismatches 437;	Indels 0;	Gaps 0;			
	Matches 3617;	Conservative							
QY	15	GGCGCACTGAGGCAAGCGT	TGAGGAGGAGGACGGCGGCAGCGCAGCGCTTCGCGGCCAG	74					
Db	260	GGAGGCGGAGGCGCGCTTGCGGTTCGCGCGCGCGCGTGTTCGAGGCTTCGCGCCGACGCGAGGCCAA	319						
QY	75	GAGCGGCGGCGCGCACTGTGTCAACGGCGCGTGGTGTCAACCGTGGCACCCCTGGCCACCGCGCAGTGCGT	134						



Db 1400 GGTCAATGTTGGTCAATCTTCCGGTCATTTGGCCCACTCTATGGATATATAAAGCGTTGTGCAGGA 1459  
Qy 1215 ATATGATTTGGATAAGAAACACATCGGAGTATATTACACACCAGGGTGCCTTTGTGCAAGG 1274  
Db 1460 ATATGATTTGGATAAGAAACACATCGGAGTATATTACATCTACAGGGTGTCTTGTGCAAGG 1519  
Qy 1275 GGGTACGGCCATAGCAGTGTATTACAGCAATAGAGCCAGGGCCCTATAGTTTCATGGTGG 1334  
Db 1520 GGGTATTTGGCCATAGTGTCTTATGATGACAGGACCAAGGCTCTGTGATGTTCAATGGTGG 1579  
Qy 1335 CTACAAGGCTTTAGTGTCCCAATAGTACCGGCTTGGAGATGATCTCTACCGATATGATGT 1394  
Db 1580 CTACAAGGCTTTAGTGTCCCAATAGTACCGGCTTGGAGATGATCTCTACCGATATGATGT 1639  
Qy 1395 GGATACCCAGATGTGGACCAATCTTTAAGGACAGCCGATTTTCCGTTTACTTGGACACAGC 1454  
Db 1640 GGATACCTCAGATGTGGACCAATCTTTAAGGACAGCCGATTTTCCGTTTACTTGGACACAGC 1699  
Qy 1455 TGTGATAGTGGAGTGGAAACCATGCTGTGTTTGGGGGAAACACACACATGACACATCTAT 1514  
Db 1700 TGTGATAGTGGAGTGGAAACCATGCTGTGTTTGGGGGAAACACACACATGACACATCTAT 1759  
Qy 1515 GAGCCATGGGCCCAATGCTTCTCTTCAGATTTCAATGGGCTATGACATTTGCGTGCACCG 1574  
Db 1760 GAGCCATGGGCCCAATGCTTCTCTTCAGATTTCAATGGGCTATGACATTTGCGTGCACCG 1819  
Qy 1575 CTGGTCAAGTGTCCAGACCATGATCTCCACCATGATGTCACAGATTTGGCCATTCAGC 1634  
Db 1820 ATGGTCAAGTGTCCAGACCATGATCTCCACCATGATGTCACAGATTTGGCCATTCAGC 1879  
Qy 1635 AGTCTTACACACAGCACCATGATGTTGTTTCCGTTGTTTCAATGATCTCTCTCTCAGCGA 1694  
Db 1880 AGTCTTACACACAGCACCATGATGTTGTTTCCGTTGTTTCAATGATCTCTCTCTCAGCGA 1939  
Qy 1695 CATCTCTGTTATTCCTCGGAAACAGTGTGATGGGCTCGAGTGAAGCGCTGTTTATG 1754  
Db 1940 CGTCTTGGCTTTTACCTCGGAGAGTGCAGTGCACACCGCAGTGAAGCTGCTTGTGTGGC 1999  
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCTGTGGGC 1814  
Db 2000 AGCAGGACCTGGTATTCGGTGTGTGGGACACACAGTGTCTCTCGATGTACCTCTGGGA 2059  
Qy 1815 GTTGGCAACTGTAGAACAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCT 1874  
Db 2060 GTTGGCAACTGTAGAACAGAAAGTTAAATCAGAAATGTTTTTCTAAAGAAACCTCT 2119  
Qy 1875 TGACCATGACAGATGTGACACAGCACACAGATTTGTTACAGCTGCACAGCCAAACCAATGA 1934  
Db 2120 TGACCATGACAGATGTGACACAGCACACAGATTTGTTACAGCTGCACAGCCAAACCAATGA 2179  
Qy 1935 CTGCCATGTGTCAATGACCAATTTGTGCCAGAAACACAGCTGTCTCAGAGGCCAGAT 1994  
Db 2180 CTGCCATGTGTCAATGACCAATTTGTGCCAGAAACACAGCTGTCTCAGAGGCCAGAT 2239  
Qy 1995 CTCCATTTTGTGTATGAGATTTGCCCCAGGATTAACCTATCTACTGTCTGTAAACAGAA 2054  
Db 2240 CTCCATTTTGTGTATGAGATTTGCCCCAGGATTAACCTATCTACTGTCTGTAAACAGAA 2299  
Qy 2055 GACCAGCTGCAGGAGTGTGCCCTTGGACCAAGAACTGCCAGTGGAGGCCCGGAATCAGGA 2114  
Db 2300 AACCAGCTGCAGGAGTGTGCCCTTGGACCAAGAACTGCCAGTGGAGGCCCGGAATCAGGA 2359  
Qy 2115 GTGCATTTGCCCTGCCAAATATCTGTGCATTTGGCTGGCATTTGGTGGAACTCATG 2174  
Db 2360 GTGCATTTGCCCTGCCAAATATCTGTGCATTTGGCTGGCATTTGGTGGAACTCATG 2419  
Qy 2175 TTTGAAATTTACTCTCCCAAGGAGATTTATGCAATGCTTAAATTTGTTCTGTAGGAACCA 2234  
Db 2420 TCTGAAATTTACTCTCCCAAGGAGATTTATGCAATGCTTAAATTTGTTCTGTAGGAACCA 2479  
Qy 2235 CAATGCCCTTTGGCTTCTCTTACACCCAGAGAGGATTAATTTGTTCTTTAAGCAGCT 2294  
Db 2480 CAATGCCCTTTGGCTTCTCTTACATCCCAAGAGAGGATTAATTTGTTCTTTAAGCAGCT 2539

Qy 2295 CGGAATATGAGTCACTCTCAGACATGTCCAAGCTCACCTTAACCCCATGGGTGCGCCT 2354  
Db 2540 TCGATTAATGCAATCATCTCAAGATATGTCCAAGCTCACTGTGACTCCATGGGTGCTCT 2599  
Qy 2355 TCGGAAGATCAATGTGTCTTACTGTCTGGGAAGATATGTCCCAATTTACAAATAGTTT 2414  
Db 2600 TCGGAAGATCAATGTGTCTTACTGTCTGGAGGATATGTCTCCATTCACAAATAGTTT 2659  
Qy 2415 ACTACAGTGGATGCGCTGTAGCCAGTGTGAGTGTCTGGAAATCTTGTGGAATTTTATCAGAAC 2474  
Db 2660 GCTGCAGTGGATGCGCTGTAGCCAGTGTGAGTGTCTGGGATCTTGTGAGAGCC 2719  
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCAACTTGCATCAACCCCACTCAATGGTGTCTGTGA 2534  
Db 2720 TAGTACTCGGGGATTAAGGCTGCAACTTGCATCAACCCCTCTCAATGGGAGCTGTGTGA 2779  
Qy 2535 AAGGCTTGCAAACCAACAGTGTAAAGCAGTGCAGACACCACTGTGCTTGGAGACAGATG 2594  
Db 2780 AAGGCTTGCAAACCAACAGTGTAAAGCAGTGCAGACACCACTGTGCTTGGAGACAGATG 2839  
Qy 2595 TGGAGATGACACAGCGGAGCTCTGAGTGCATGTGTGCTGAGCAACATGAGCAGTGTGT 2654  
Db 2840 TGGGAGTGCATAGCAGCAGCTCGAGTGTGATGTGTGAGTAAACATGAGCAGTGTGT 2899  
Qy 2655 GGAATCCAAATGCTTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714  
Db 2900 GGAATCCAAATGCTTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2959  
Qy 2715 GAGCAGCTGCCCTCGAAATTTGTTGAGGCTACTGTACCTGTAGTCAATTTGCTTGGAGCA 2774  
Db 2960 GAGCAGCTGCCCTCGAAATTTGTTGAGGCTACTGTACCTGTAGTCAATTTGCTTGGAGCA 3019  
Qy 2775 ACCAGGCTGTGGCTGTACTGTATCCAGCAATCTGGCAAAAGGAAATGCAATGAGGG 2834  
Db 3020 GCGAGGCTGTGGTGTGTGTACTGTATCTAGCAATCTGGGAAGGAAATGTTATGAGGG 3079  
Qy 2835 TTCTTATAAGGACAGTGAAGATGCTTTCGAAGCCCTTACAGGAAATTTCTATCCACA 2894  
Db 3080 CAGCTATAAGGACAGTGTGAAGATGCGCTTCAAGCCCTCTGCAGGAAATGTATCCACA 3139  
Qy 2895 GCGCTGTCTCAATTTCCAGCATGTCTTAGAGGACAGCAGATACCACTGGTCTTTCATTC 2954  
Db 3140 GCGCTTCTGAACTCCAGCATGTCTTAGAGGACAGCAGATACCACTGGTCTTTCATTC 3199  
Qy 2955 CTGTCCAGCTTGCCTTGCACCGCCACAGTAAATGCAATCAGAGCATCTGTGTAGAA 3014  
Db 3200 CTGTCCAGCTTGCCTTGCACCGCCACAGTAAATGCAATCAGAGCATCTGTGTAGAA 3259  
Qy 3015 GTGTGAGAACCTTGACCCACAGCAGCAGTGTCTAGAGGACAGCAGATACCACTGGTCTTACGGTGA 3074  
Db 3260 GTGTGAGAACCTTGACCCACAGCAGCAGTGTCTAGAGGACAGCAGATACCACTGGTCTTACGGTGA 3319  
Qy 3075 TCCCAACCAATGGAGGAAATGTACAGCCTTGAAGTGAATGGGCAACCGCTCTCTGTGCAA 3134  
Db 3320 CCGGACTTATGGAGGCAATGTACAGCCATGCAAGTGAATGGGCAACCGCTCTCTGTGCAA 3379  
Qy 3135 CACCAACAGGCGCAAGTGTCTTGCACCAACAGGCGGTCAAGGGGACAGTGCAGCT 3194  
Db 3380 CACCAACAGGCGCAAGTGTCTTGTACCAACAGGAGTGTCAAGGGGACAGTGCAGCT 3439  
Qy 3195 ATGTGAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACTGTTTATTTACTCT 3254  
Db 3440 ATGTGAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACTGTTTATTTACTCT 3499  
Qy 3255 TCTTATTTGACATCAGTTTACCTTTAGTCTATCCAGGAGATGATCGCTTATTTACAGC 3314  
Db 3500 TCTTATTTGACATCAGTTTACCTTTAGCTGTCTCCAGGAGACAGCGCTACTTACAGC 3559  
Qy 3315 TATCAATTTTGTGGCTACTCTGTGACGAAACAAACAGGGATTTGGACATGTTTCAATGTC 3374  
Db 3560 CATCAATTTTGTGGCTACTCTGTGATGAACAAACAGGGATTTTGCATGTTTCAATGTC 3619

QY 3375 CTCAAGATTTCAACCTCAACATCACTCGGCTGCCAGTTTCTCAGCTGGAAACCCAGGC 3434  
 Db 3620 CTCACAAAATTTCAACCTCAACATCACTCGGCTGCCAGTTTCTCAGCTGGAAACCCAGGC 3679  
 QY 3435 TGAAGAAGAGATGCTGCTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTC 3494  
 Db 3680 TGAAGAAGAGATGCTGCTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGTTTCTC 3739  
 QY 3495 TAATGAGAAGTTTGAATTTTCGCAACCAACCAATATACATTTCTTTGTTTATGTCAGTAA 3554  
 Db 3740 TAATGAGAATTTGAATTTTCGCAACCAACCAATATACATTTCTTTGTTTATGTCAGTAA 3799  
 QY 3555 TTTCACCTGCGCCATCAAAATTCAGATGCTTCTCAGCAAGCAAGCAATTTTATGACCT 3614  
 Db 3800 TTTCACCTGCGCCATCAAAATTCAGATGCTTCTCAGCAAGCAAGCAATTTTATGACCT 3859  
 QY 3615 GGTACAGTCTTCTGTCATTTCTTCAAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3674  
 Db 3860 GGTACAGTCTTCTGTCATTTCTTCAAGTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3919  
 QY 3675 GGTGGAAGATCAAAAGTTTGTGGGCTCCAGACGTAGAGCAACTTCTTCGAGA 3734  
 Db 3920 GGTCTGGAAGATCAAGCAGAGCTGTGGGCTCCAGACGTAGAGCAACTTCTTCGAGA 3979  
 QY 3735 GATGCAACAGATGGCCAGCGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3794  
 Db 3980 GATGCAACAGATGGCCAGCGCTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4039  
 QY 3795 GGAAGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTCCCAAAACCCATTCGACTGA 3854  
 Db 4040 AGAAGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTCCCAAAACCCATTCGACTGA 4099  
 QY 3855 GCGGTGTTTGGCAACAAAGCCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3914  
 Db 4100 GCGGTGTTTGGTAAACAAAGCCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4159  
 QY 3915 GGTGGGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3974  
 Db 4160 GGAAGGATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4219  
 QY 3975 TTCTCAGCAGATGCCGATGTTACAGGAGAGTACAGGAGCGCTGAGAAACCGGAAGCA 4034  
 Db 4220 TTCTCAGCAGATGCCGATGTTACAGGAGAGTACAGGAGCGCTGAGAAACCGGAAGCA 4279  
 QY 4035 GCAGCCCTCTGCACAGCTGGACCTGCATCTGA 4068  
 Db 4280 GCAGCCCTCTGCACAGCTGGACCTGCATCTGA 4313

RESULT 9  
 AB038387 Rattus norvegicus atrn mRNA for attractin, complete cds.  
 LOCUS AB038387 8739 bp mRNA linear ROD 18-JAN-2001  
 DEFINITION Rattus norvegicus atrn mRNA for attractin, complete cds.  
 ACCESSION AB038387  
 VERSION AB038387.1 GI:12275307  
 KEYWORDS attractin.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 1 (sites)  
 Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T.,  
 Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G. S., Nagao, M., Ushijima, T.  
 and Serikawa, T.  
 TITLE Attraction/mahogany/zitter plays a critical role in myelination of  
 the central nervous system  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)  
 MEDLINE 21143347  
 PUBMED 11209055  
 REFERENCE 2 (bases 1 to 8739)  
 Kuramoto, T., Serikawa, T. and Ushijima, T.  
 Direct Submission  
 TITLE

JOURNAL Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center  
 Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,  
 Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.go.jp,  
 Tel:81-33542-2511 (ex.4521), Fax:81-35565-1753)  
 Location/Qualifiers  
 1..8739  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /strain="WTC"  
 /db\_xref="taxon:10116"  
 1..8739  
 /gene="atrn"  
 61..4359  
 /gene="atrn"  
 /codon\_start=1  
 /product="attractin"  
 /protein\_id="BAB21017.1"  
 /db\_xref="GI:12275308"  
 /translation="MVAAAATAEATRLRGYTTTATAAPAGWKSRQHRPCATGAWRP  
 WPRAGLCLPRVLSRALSPPLLPILLPLLSLLPLPREAEEAAVAASGSAARAK  
 ECDPCVNGRCNPGTCQVCTGWBGEQOHCGRFRLTGSSGFVTDGPNYKYTK  
 CTWLEGPNKIMRLRFNFHATEGSDWHLVYVDGSIYAPLIAFSGLIVPERDNET  
 APEVTVSYALLHFFDAAYNLTFGNITYNFDMNPNCSGRGCKSSNSTVECEC  
 SENMKGSCDIPHCCTDNCGPFRGICNASDTRGCGSCPWMQGPCSIPVPAQSFWR  
 REYSDKLPRASHKAVNGNIMVVGGMENHSYSMLAYDLASREMLSNHSVNSV  
 VVRGHSALHKKIKYMGKIDSTGNVNEHRLVFIHNSWVLLTPAKQDQYAVGH  
 SAHIVTSSGRVVMVIFGHCPLYGISVVOEYDLEKNTNLSILQTQCALVGGYGHSS  
 VYHRTKALYVHGKYKAPSAANKYRLADLYRYHYDTQWTILKDSRFYRHLTAIVS  
 CTMLVFGNTNDTSMHGAKCFSSDFMAYDIADRWSLPRPELHLDVNRFGHSVL  
 HNTWVYFGFNSLLLSLVFTSEQCDAAHSEAAACVAGPGICLMDTQSRCTSWL  
 LATQEAELKSECFKRTLDHCDQHTDCYCTANTDNCWDCNCPVNHSCTEG  
 QISAKYDNCPKDNPYCNKKTSCRSALDQNCWEPRQECIALPENIGIWHLV  
 GNSCLITAKENDYAKNSCRHNFLASLSQKVEFLVKQLRLNQSSOSTSKLTL  
 TFWGLRKINVSYWCEDMSPFTNSLLQWMPSEPSDAGFCGILSEPSTRGKAATCIN  
 PLNGSVRERPAHSAKQCTPCALRTACGECTSSSECSMNCWQCVNAYASFP  
 PQGMEWTYSSCPGSGYCTGCHLEQFCGCGCTDPSNTGKGTIEGSKYGPVKM  
 PSHASTGNVYPOPLNNSMCLDSRKNXNSFHCPCACQSHGKINOSICRQBDLTT  
 GKHCETCSGFVGPDTNGKQCPCKXNGHSLCNTNTGKCFCTTVKGYGECOLCEVE  
 NRYQGNPLKGTCTYLLIDYQFTSLQEDDRYYTAINFVATPEQNRDLDFINASK  
 NFNLNIWATSPFAGTQIGEEVPVVKINIKKYSFNSKFKDFRNHNITTFYVSN  
 FTWPIKIQIAFSQHSNFMVLOVFTFSCFLSLLVAAVVKIKPCQKWAARREQLL  
 REMQMASRPFASYNVALETDEPDLIGSITKVPKPIALBPCKGNKAAVLVSVRL  
 PRGLGGIPPGQGLAVASALVDISQMPIVYKESGAVRNRRKQPPAQPGTCI"

ORIGIN  
 Query Match 82.4%; Score 3353.2; DB 10; Length 8739;  
 Best Local Similarity 89.2%; Pred. No. 0;  
 Matches 3616; Conservative 0; Mismatches 438; Indels 0; Gaps 0;  
 QY 15 GCGCGCAACTGAGGCAAGCTGAGGAGGAGGACGCGCGCGCGCTCGCGGCGAG 74  
 Db 306 GGAGGCCGAGCCCGTGCCTGGCGCGCTGTGTTCAGGCTCGCGCGAGCTGAGGCCAA 365  
 QY 75 GAGCGCGCGCGCGCTGTGTCAACGCGCGCTGTGTCAACCGCTGCGACCGCGCGAGTGGCT 134  
 Db 366 GGAATGTGACCGCGCGCTGTGTCAACGCGCGCGCTGTGTCAACCGCGCGAGTGGCT 425  
 QY 135 CTGCCCCCGCGCGCTGGGTGGCGAGCAATGCCAGCTCGCGCGCGCGCTTCAGACTAAC 194  
 Db 426 CTGCCCCCGCGCGCTGGGTGGCGAGCAATGCCAGCTCGCGCGCGCGCTTCAGACTAAC 485  
 QY 195 TGGATCTTCTGGGTTTGTGACAGATGGACCTTGGAAATTTATAAATACAAAGTGCAC 254  
 Db 486 TGGATCTTCTGGATTTGTAAACAGATGGACCTTGGAAATTTATAAATACAAAGTGCAC 545  
 QY 255 GTGGCTCAATTGAGCAGACGCAATAGATAATGAGACTGTGTTTCAATCATTTTGTCTAC 314  
 Db 546 ATGGCTCAATTGAGCAGACGCAATAGATAATGAGACTGTGTTTCAATCATTTTGTCTAC 605  
 QY 315 AGAGTGTAGTTGGGACCACTTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGT 374  
 Db 606 TGAATGTAGCTGGGACCACTTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGT 665

QY 375 TGCTGCAATTTAGTGGCCCTCATTTGTTCTCTGAGAGAGATGGCAATCAGAGCTGTCTCCTGAGGT 434  
DB 666 TGCTGCCCTTTAGTGGCCCTCATTTGTTCTCTGAAAGAGATGGCAATGAGAGCGGCCCTCAAGT 725  
QY 435 TGTGGCCACATCAGGTTATGCTTGTGCTGCAATTTTTTTAGTGTGCTGCTTATTAATTTGAC 494  
DB 726 CACTGTCACTTTCAGGTTATGCACTACTACTGCAATTTTTTTCAAGTGTGCTGCTTATTAATCTGAC 785  
QY 495 TGGATTTAATATTACTTTACAGTTTGTATGTGTCCAAATAACTGTCCAGGCCGAGAGA 554  
DB 786 TGGATTTAATATTACTTTACAAATTTGACATGTGCGCAATTAATGTCTCAGGCCGAGAGA 845  
QY 555 GTGTAAGATCAGTAAATAGCAGCAGATCTGTGTGAATGTGAATTTCTGAAACTGTGAAAGG 614  
DB 846 ATGTAAGAGCAGTAACAGCAGCAGCAGTGTGTGAGTGTGAATGTTCTGAAACTGTGAAAGG 905  
QY 615 TGAAGCATGTGACATTTCTCACTGTACAGACAATGTGTGTTTCTCTCATCGAGSCATCTG 674  
DB 906 GGAGTCTGTGACATTTCTCACTGTACAGACAATGTGTGTTTCTCTCATCGAGSCATCTG 965  
QY 675 CAATTTCAAGTGTATGTACAGAGATGCTCTCTCTTCTCAGACTGCGAGGTCCTGGATGTT 734  
DB 966 TACGCAAGTGTACACAGAGGCTGCTCTCTCTCTCTCACTGCGAGGTCCTGGATGTT 1025  
QY 735 AGTTCTGTGTACCACTTAACAGTCAATTTGAGCTCGAGAGATATTCTTAACCTTAAGCT 794  
DB 1026 AATTCCTGTGCCAGCTTAACAGTCTTTTGGACTCGAGAGATATTCTTAACCTTAAGCT 1085  
QY 795 CCCCAGAGCATCTCATAAAGCTGTGTCAATGGAACATTTATGTGGTGTGTTGGAGGATA 854  
DB 1086 TCCCAGAGCCTCTCATAAAGCTGTGTCAATGGAATATTATGTGGTGTGTTGGCGGTA 1145  
QY 855 TATGTTCAACACTCAGATTAACATGTTCTAGCGTATGACCTTGCTTCTAGGAGTG 914  
DB 1146 TATGTTCAACACTCAGATTAACAGTGTGTTTGTAGCGTATGACCTGCTTCTAGGAGTG 1205  
QY 915 GCTTCCACTAAACCGTCTGTGAACATGTGTTGTAGATATGGTCACTTTTGGCAAT 974  
DB 1206 GCTTTCCTAAACCTGTGTGAACATGTGTTGTAGATATGGTCACTTTTGGCAAT 1265  
QY 975 ATACAGGATTAATTTACATGTATGAGGAGAAATTTGATTTCACTGGGAATGTGACCAA 1034  
DB 1266 ACATAAGGATAAATCTACATGTATGAGGAGAAATTTGATTTCAACAGGGAATGTGACCAA 1325  
QY 1035 TGAGTTGAGAGTTTTTACATTCATATGAGTCACTGGTGTGTTGACCCCTTAAGCATA 1094  
DB 1326 CGAGCTGAGAGTATTTTATATTTCAATGAATGATGGTATTTGTTAACTTCAAGGCTAA 1385  
QY 1095 GGAGCAGTATGCACTGTTGGGCACTCTGCACACATTTGTATCACTGAAAGATGGCCGAGT 1154  
DB 1386 GGATCAGTATGCACTGTTGGGCACTCTGCACACATTTGTATCACTGTCATCTGGCCGTGT 1445  
QY 1155 GGTCAATGCTGTTGATCTTTGGTCTACTGCTCTCTATGAGATATTAAGCAATGTGACGAA 1214  
DB 1446 GGTCAATGCTGTTGATCTTTGGTCTACTGCTCTCTATGAGATATTAAGCAATGTGACGAA 1505  
QY 1215 ATATGATTTGATTAAGAACACATGGAGTATATTACACACCCAGGTCCTTGTGCAAGG 1274  
DB 1506 ATATGATTTGATTAAGAACACATGGAGTATATTACAACTCAAGTGTCTAGTGAAGG 1565  
QY 1275 GGGTTACGGCATTAGAGTGTGTTACGACCATAGGACAGGGCCCTATACGTTATGGTGG 1334  
DB 1566 GGGTTATGGCCACAGCAGTGTGTTATGATCAAGGACCAAGGCTCTGTATGTTTATGGTGG 1625  
QY 1335 CTACAGGCTTTCAAGTCCAAATAGTACCGGCTTGCAGATGATCTCTACCGCATATGATGT 1394  
DB 1626 CTACAGGCTTTCAAGTCCAAATAGTACCGGCTTGCAGATGATCTCTACAGATACCATGT 1685  
QY 1395 GGATACCCAGATGTGGAACCATTTCTTAAGGACAGCCGATTTTTCCGTTATCTTGACACAGC 1454  
DB 1686 GGACACTCAGATGTGGAACCATTTCTTAAGGACAGCCGATTTTTCCGTTATCTTGATACAGC 1745  
QY 1455 TGTGATGATGAGTGGACCATGCTGTTTGGGGGAAACACACACATGACATCTAT 1514

DB 1746 TGTGATGATGAGTGGAAACCATGCTGGTGTTCGGAGGAAACACACAAATGACACTTCCAT 1805  
QY 1515 GAGCCATCGCCCAAAATGCTTCTCTCAGATTTTCATGSCCTTATGACATTTGCTGTGACCG 1574  
DB 1806 GAGCCACCGTGCCTAAATGCTTCTCTCAGACTTCTCATGGCTTATGACATTTGCTGTGACCG 1865  
QY 1575 CTGCTCAGTGTCTCCAGACCTGATCTCCACCATGTGTCAACAGATTTGGCCATTCAGC 1634  
DB 1866 ATGCTCAGTGTCTCCCCAGACCTGAGCTCCATCAGATGTCAACCGATTTGGCCACTCAGC 1925  
QY 1635 AGTCTTACACACAGCACCATGTATGTGTTGCTGGTGTTCATATGATCTCTCTCTCAGGA 1694  
DB 1926 AGTCTTGCACACAGCACCATGTATGTGTTGCGGGCTTCAACAGCTCTCTCTCTCAGTGA 1985  
QY 1695 CATCTGTGATTTCACTCTCGAAACAGTGTGATGCTCGAGTGAAGCGGCTGTTTTCAGC 1754  
DB 1986 CGTCTTGTCTTCACTCTCGAGCAGTGTGATGCAACCGCAGCAAGCTGTCTGTGGC 2045  
QY 1755 AGCAGGACTCGTGTATTCGTTGTGTGGAACACAGGGTCTCTCAGTGTATCTCTGTGGC 1814  
DB 2046 AGCAGGACTCGGCTTCCGCTGTCTGTGGGACACACAGTGTCTCGATGTACTCTCTGGGA 2105  
QY 1815 GCTGCAACTCATGAACCAAGAAAGTAAATCAAGATGTTTTCCTCAAAAGAACTCT 1874  
DB 2106 GTTGGCACTGAGNAACAGCAGAAAGTTAAATCAGATGTTTCTCTAAAGAAACCT 2165  
QY 1875 TGACCATGACAGATGTGACAGCAGCAGATGTTTACAGTGTGACAGCAGCAGCAATGA 1934  
DB 2166 TGACCATGACAGATGTGACAGCAGCAGATGTTTACAGTGTGACAGCAGCAGCAATGA 2225  
QY 1935 CTGCCACTGTGCAATGACCAATGTGTCCTCCAGCAACACAGCTGCTCAGAAAGCCAGAT 1994  
DB 2226 CTGCCACTGTGCAATGACCAATGTGTCCTCCAGCAACACAGCTGCTCAGAAAGCCAGAT 2285  
QY 1995 CTCCATTTTGTAGTATGAGAAATGCCCCAAGGATAACCTATGTACTGTATTAACAGAA 2054  
DB 2286 CTCCATTTGCAAGTATGCAATTTGCCCAAGGATAACCTATGTACTGTATTAACAGAA 2345  
QY 2055 GACAGCTGTGAGGAGTGTGCTCTGACAGAACTGCGAGGAGGAGCCCGGATCAGGA 2114  
DB 2346 AACAGCTGTGAGGAGTGTGCTCTAGATCAGAAATGCGAGTGGGAACTCGAAATCAGA 2405  
QY 2115 GTGCATTTGCTGCCCCGAAATATCTGTGCAATTTGGTGTGCTGTTGTTGAACTCATG 2174  
DB 2406 GTGCATTTGCTGCCCCGAAATATCTGTGCAATTTGGTGTGCTGTTGTTGTTGTTG 2465  
QY 2175 TTTGAAATTTACTACTGCTCAAGGAGATTTATGCAATGCTTAAATGTTCTGTAGGAACCA 2234  
DB 2466 TCTGAAATTTACTACTGCTCAAGGAGATTTATGCAATGCTTAAATGTTCTGTAGGAACCA 2595  
QY 2235 CAATGCCCTTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGCTCTTAAGCAGCT 2294  
DB 2526 CAATGCCCTTTTGGCTTCTCTTACGTTACGTTCCAGAAAGGTGAGTTTGTCTCTTAAGCAGCT 2585  
QY 2295 GCGAATATGCACTCATCTCAGAGATGTCCAAAGTCACTTCACTTACCCCACTGGGTGCGCT 2354  
DB 2586 TCGAATTAATGCAATCATCTCAGAGCAGTCCCAAGCTCACTGCTGCTCACTGGTGTGCT 2645  
QY 2355 TCGGAAGATCAATGTGCTCTACTGCTGCGAAGATATGTCCTCAATTTTACAAATAGTGT 2414  
DB 2646 TCGGAAGATCAATGTGCTCTACTGCTGCGAAGATATGTCCTCAATTTTACAAATAGTGT 2705  
QY 2415 ACTACAGTGTGCGGTCTGAGCCAGTGTGCTGAGTATGCTGTGGAATTTTATCAGAAC 2474  
DB 2706 GCTGCACTGATGCTGCTGAGCCAGTGTGCTGAGTGTGCTGCTGCTGCTGCTGCTGCT 2765  
QY 2475 CAGTACTCGGGAGTGTGAGGCTGCAACCTGCTCAACCCACTCAATGCTGTGTGTGTA 2534  
DB 2766 TAGTACTCGGGAGTGTGAGGCTGCAACCTGCTCAACCCACTCAATGCTGTGTGTGTA 2825  
QY 2535 AAGGCTCTCAAAACCACTGCTGAAGCAGTGTGCGGACCACTGTGCTGCTGAGGACAGATG 2594



917	TGAGGTCACTGTGCATCTTACAGGTTATGCACTGCTGCATTTTTTTTCAGTGAATGCTGCTTATAA	976
489	TTTGACTGGATTTTAATATTAATCTTACAGTGTGTTGATATATGTTCCAAATAACTGCTCAGCGCG	548
977	TCTGACTGGATTTTAATATCACTTAACAATTTTGACATGTTCCGGAATAAATTTGCTCAGCGCG	1036
549	AGGAGAGTGTAAGATCAGTAATACAGCGGATACGTGTTGAATGTGCAATGTTCTGAAAACGTG	608
1037	AGGAGAGTGTAAGAGCAGTAACAGCAGCGCTGTTGAGTGTGAAATGTTCTCTGAAAACGTG	1096
609	GAAAGGTCAAGCATGTGCACATTCCTCACTGTACAGACAACTGTGTGGTTTTTCTCATTCGAGG	668
1097	GAAAGGGGAGTTCGTGTGCACATTCCTCACTGTACAGACAACTGTGTGGCTTTCTCATCCGAGG	1156
669	CATCTGCAATTCAGATGATGTACAGAGATGCTCTCTGCTTCTCAGACATGGCAGGGTCTCTGG	728
1157	CATCTGTAATGCAAGCGATACCAAGAGGGTGTCTCTGCTTTTCTCACTGGCAGGGTCTCTGG	1216
729	ATGTTTCAGTTCCCTGTACAGCTAACCACTGATTTTGGACTCGAGAGGAATATTTCAACTT	788
1217	ATGTTCAATTCCTGTGCCAGCTAACCACTTTTTTGGACTCGAGAGGAATATTTCTGATTT	1276
789	AAAGCTCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAAAAATTAATGTGGGTGTGTTGG	848
1277	AAAGCTTTCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAAAATATAATGTGGGTGTGTTGG	1336
849	AGGATATATGTTCAACCACTCAGATTTATAACATGGTTCTAGCGTATCACTTGTCTTCTAG	908
1337	CGGATATATGTTTCAACCATTCAGATTTACAGCATGGTTTTTACGGTATGACCTGACTTCTAG	1396
909	GGAGTGGCTTCCCATAAACCGTTCCTGTGAACAAATGTGGTTGTTAGATATGGTCAATCTTTT	968
1397	GGAAATGCTTCCCATAAACCATTCCTGTGAACAGTGTGTTGTAAGATATGGTCAATCTTTT	1456
969	GGCATTTATACAGGATAAATTTACATGTATGGAGGAAAAATTTGATTTCAACTGGGAATGT	1028
1457	GGCATTTACATAGGATAAATCTACATGTATGGAGGAAAAATTTGATTTCAACAGGGAACGT	1516
1029	GACCAATGAGTTGAGAGTTTTTTACATTCATTAATGAGTCAATGGGTGTGTTGACCCCTAA	1088
1517	GACCAATGAGCTGAGAGTATTTTCATATTCATATGANTCATGGGTATGTTTAATCCGAA	1576
1089	GGCAAAAGGAGCATGATGACGTGTGTTGGGCACTCTGCAACATTTGTTTACAGGAATGG	1148
1577	AGCTTAAGGATCAGTATGACGTGTGTGGCACTCTAGCAACATTTGTTTACACTGGCATCTGG	1636
1149	CCGAGTGTCTCATGCTGTGTCACTTTTGTGTCACTGCCCTCTCTATGGATATATAAGCAATGT	1208
1637	CCGTGTGGTCAATGTTGGTTCATCTTCGGTTCATTTGCCCACTCTATGGATATATAAGCGTTGT	1696
1209	GCAGGAATATCATTTGGATAAGAAACATCGAGTATATTAACACCCAGGGTGCCCTTGT	1268
1697	GCAGGAATATGACTTTGGAAGAAACACATCGAGTATATTAATCACTCAGGGTGCTCTTGT	1756
1269	GCAAGGGGGTTACGCCCATAGCATGTGTTTACGCCATAGGACCAAGGCCCTTACAGTTCA	1328
1757	GCAAGGGGGTTATGGCCACAGTAGTGTGTTATGATGACAGGACCAAGGCTCTGTACAGTTCA	1816
1329	TGTTGGCTACAGAGCTTTTCAGTGCCTAATAAGTACCGCGTTTGCAGATGATCTCTACCGATA	1388
1817	TGTTGGCTACAGAGCTTTTCAGCGCCCAACAAATACCGGCTTGCAGATGACCTTACAGATA	1876
1389	TGATGTGGATACCCAGATGTGGACATTTCTTAAGGACAGCGGATTTTTTCGTTACTTTGCA	1448
1877	CGATGTGGATACCTCAGATGTGGACCATTTCTTAAGGACAGCGGATTTTTTCGTTACTTTGCA	1936
1449	CACAGCTGTGATAGTGTAGTGAACCATGTGGTGTGTTTTGGGGAAAAACACACAAATGACAC	1508
1937	TACAGCTGTGATGTAGTGTGAACCATGCTGTGTGTTTGGAGGGAACACACACATGACAC	1996
1509	ATCTATGAGCGAATGGCGCCAAATGCTTCTCTTTAGATTTTCATGGCCTATGACATTTGCTGT	1568
1997	TTCCATGAGCGCACGGTGCCTAAATGCTTCTCTCGAATTCATGGCTTATGACATTTGCTGT	2056





VVRYGHS LALHDKKIYMGKIDSTGNVTNELRVFHHNESWLLTPKAKDQYAVVGH  
SAHIVTLSSGRVVMVLVI FGHCPYIGYISVVQVEYDLEKNTWSILQ7GALVQCGYGHSS  
VYDHRTKALVYHGGYKAFANKYRLADLYRVHVD7OMWT7ILKDRFFRYLHTAVIS  
GTMVLFGGTHNDTSVSHGAKCFSSDPMAVD7IACDEWSVLP3PELHHDVNRFGHSAVL  
HNSTWVFGFNLSLL7DLV7F7SEQCDHRSEACVAAGPGRICLWDTQ7SRSCT7SWE  
LATEQREKLS7CF7SK7RL7DHR7CD7H7D7C7S7TANT7ND7C7W7D7H7C7V7P7VNH7S7TEG  
Q7S7I7A7K7Y7D7N7C7P7D7N7P7Y7N7K7T7S7C7R7S7C7A7L7D7Q7W7E7P7R7Q7E7C7I7A7L7P7E7N7I7G7H7L7V  
GNSCL7K7T7T7A7K7E7N7D7A7K7S7C7R7N7H7A7F7L7A7S7T7Q7K7E7F7L7Q7L7R7M7S7Q7S7K7L7T  
TPW7G7L7K7N7V7W7C7H7E7D7M7S7P7T7N7S7L7L7O7M7P7S7E7P7S7D7A7G7F7G7I7S7P7S7R7G7L7A7A7C7I7N  
PLNG7S7C7E7R7P7A7N7S7A7K7O7C7P7C7A7L7R7T7A7C7B7E7T7S7S7E7C7W7C7M7S7N7Q7C7V7D7S7N7A7Y7A7F7P  
F7G7O7M7W7Y7M7S7S7P7P7E7N7C7S7Y7C7T7S7C7H7L7B7Q7P7G7C7M7C7D7P7S7N7T7G7K7C7I7E7G7S7K7E7G7P7V7M  
P7S7H7A7S7T7N7V7P7O7L7L7N7S7C7L7E7S7R7Y7N7S7F7I7H7C7P7A7C7Q7C7H7S7K7C7I7N7Q7S7I7C7E7C7E7D7L7T  
G7K7E7C7I7S7Y7G7D7P7N7G7K7C7Q7P7K7C7N7G7H7A7S7I7N7T7G7K7F7C7T7I7G7V7K7E7B7Q7C7L7E7V7E  
N7R7Y7Q7N7P7L7K7G7T7C7Y7T7L7L7I7D7Y7F7T7S7L7Q7D7R7Y7T7A7I7N7F7A7T7P7D7B7Q7N7R7D7L7M7F7I7N7A7S7K  
N7F7N7L7I7T7W7A7T7S7F7P7A7G7T7Q7B7E7V7P7W7S7K7T7N7I7K7E7Y7K7D7S7F7N7E7K7P7D7F7R7N7H7P7I7T7F7V7S7N  
F7T7W7P7I7K7I7Q7V7R7T7S7"

## ORIGIN

Query Match	72.08;	Score 2928.8;	DB 10;	Length 4500;
Best Local Similarity	88.94;	Pred. No. 0;		
Matches 3167;	Conservative 0;	Mismatches 397;	Indels 0;	Gaps 0;
QY	15	GGCGGCAACTGAGCAAGGCTGAGGAGGAGGACGGCGGGGACGGCAGCGCTCGCGGGCAG	74	
Db	306	GGAGGCCGAGGCCCTCGCGTGGCGGCGGTGTGTAGGCTCGGCCCGAGCTGAGGCCAA	365	
QY	75	GAGCGGCGGCGCGCACTGTGTCAACGGCGGTGCTGTCAACCCCTGTGGCAACCGGCGAGTGGCT	134	
Db	366	GGAATGTGACGGCGCGGTGTCAACGGCGCGCTGTCAACCCCTGGCACCGGGCAGTGGCT	425	
QY	135	CTGCCCGCGGCTGGGTGGCGGAGCAATGCCAGCACTGGGGGGCGGCTTCAGACTAAC	194	
Db	426	CTGCCCGCACGGGCTGGGTGGCGGAGCAATGCCAGCACTGGGGGGCGGCTTCAGACTAAC	485	
QY	195	TGATCTTCTGGGTTCTGTACAGATGGACCTGTGAATTTATAAATACAAAACGAAGTGCAC	254	
Db	486	TGATCTTCTGGATTTGTAAAGATGGACCTGGGAATTTAATATATAACGAATATGCAC	545	
QY	255	GTGGCTCATTTGAAGGACGAGCAATAGATAAATGAGACTTCGTTTCAATCATTTTGTCTAC	314	
Db	546	ATGGCTCATTTGAAGGACGAGCAATAGATAAATGAGACTTCGATTCAAACCATTTTGTCTAC	605	
QY	315	AGAGTGTAGTTGGGACCAATTTATGTTTATGATGGGACTCAATTTATGCAACGGCTAGT	374	
Db	606	TGAATGTAGCTGGGACCAATTTATGTTTATGATGGGACTCAATTTTACGCACCTCTGAT	665	
QY	375	TGCTGCATTTAGTGGCTCATTTGTTCTGTAGAGAGATGGCAATGAGACTGTCCCTGAGGT	434	
Db	666	TGTCGCTTTTAGTGGCTCATTTGTTCTGAAAGAGATGGCAATGAGACGGCCCTGAAGT	725	
QY	435	TGTTGCCACATCAGGTTATGCTTGTGCAATTTTTTTTGTGTGATGCTGCTTTAATTTGTAC	494	
Db	726	CACGTGCACCTCAGGTTATGCACTACTGCAATTTTTTTTTCAGTGATGCTGCTTTAATCTGAC	785	
QY	495	TGATTTAATAATTAATTAAGTTTGTGATGTGTCCAAATAACTGTCTCAGCGCGGAGAGA	554	
Db	786	TGATTTAATAATTAATTAAGTTTGTGATGTGTCCCGGAATAATGTCTCAGCGCGAGAGA	845	
QY	555	GTGTAGATCAGTAATAGCAGGATACTGTTGTAATGTGAATGTTCTGAAAACCTGGAAAGG	614	
Db	846	ATGTAAGAGCAGTAAACAGCAGCAGCACTGTTAGTGTGAATGTTCTGAAAACCTGGAAAGG	905	
QY	615	TGAAGCAATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGGCATCTG	674	
Db	906	GGAGTCTGTGACATTCCTCACTGTACAGACAACTGTGGCTTTCCTCACCGAGGCATCTG	965	
QY	675	CAATTCAGATGATGTGAGGAGTGTCTCTGCTTCTCAGACTGGCAGGGTCCTGGATGTTTC	734	
Db	966	TAAAGCAAGTGTACACAGAGGGTGTCTCTGCTTCCCTCACTGGCAGGGTCTCTGGATGTTTC	1025	
QY	735	AGTTCTGTACAGCTAACCAAGTCAATTTTGGACTCGAGAGGAATTTCTAACTTTAAGCT	794	
Db	1026	AATTCCTGTGCACTAACCAAGTCTTTTGGACTCGAGAGGAATTTCTCACTTTAAGCT	1085	

QY	795	CCCAGAGCATCTCTATAAGCTGTGGTCAATGGAAAAATTAATGTGGGTCTTGTGGAGGATA	854						
DB	1086	TCCAGAGCCTCTCTATAAGCTGTGGTCAATGGAAAAATTAATGTGGGTCTTGTGGCGGTA	1145						
QY	855	TATGTTCAACCATCAGATTATAACATGTTCTAGCGTATGACCTTCTCTAGGAGTGTG	914						
DB	1146	TATGTTCAACCATCAGATTATAACATGTTCTAGCGTATGACCTTCTCTAGGAGTGTG	1205						
QY	915	GCTTCCACTAAACCGTCTCTGGAACAATGTTGTTAGATATGATGTTCTTCTTGGCATT	974						
DB	1206	GCTTCCACTAAACCATCTCTGGAACAATGTTGTTAGATATGATGTTCTTCTTGGCATT	1265						
QY	975	ATACAAGATATAAATTTACATGTTATGAGGAAAAATTTGATTCACTGGGATGTGACCAA	1034						
DB	1266	ACATAAGGATATAAATTTACATGTTATGAGGAAAAATTTGATTCAACAGGGAATGTGACCAA	1325						
QY	1035	TGAGTTGAGAGTCTTTCACATTTCAATATGATGTTGTTGACCTTACCCCTTAAGGCAAA	1094						
DB	1326	CGAGCTGAGAGTATTTCAATTTCAATATGATGTTGTTGTTTAAACCCCAAGCTAA	1385						
QY	1095	GGAGCAGTATGACGTGTGGGCACTCTGCACACATTTGTTACACTGAAGAAATGGCCGAGT	1154						
DB	1386	GGATCAGTATGACGTGTGGGCACTCTGCACACATTTGTTACACTGTCATCTGCGCGTGT	1445						
QY	1155	GGTCATGCTGTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGA	1214						
DB	1446	GGTCATGTTGGTCACTTTGGTCACTTTGGTCACTTTGGTCACTTTGGTCACTTTGGTCA	1505						
QY	1215	ATATGATTTGGATTAAGAACACATGGAGTATTTACACACCCAGGGTGCCTTGTGCAAG	1274						
DB	1506	ATATGATTTGGATTAAGAACACATGGAGTATTTACACACCCAGGGTGCCTTGTGCAAG	1565						
QY	1275	GGTTTCGGCCATAGCAGTGTTTACGACCATAGGACCGAGGCGCTATACGTTTCATGTTG	1334						
DB	1566	GGTTTCGGCCATAGCAGTGTTTACGACCATAGGACCGAGGCGCTATACGTTTCATGTTG	1625						
QY	1335	CTACAGGCTTTTCAAGTGCATTAAGTACCGGCTTTCGACGATGATCTCTACCGGATATGAT	1394						
DB	1626	CTACAGGCTTTTCAAGTGCATTAAGTACCGGCTTTCGACGATGATCTCTACCGGATATGAT	1685						
QY	1395	GGATACCCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTTCGTTTACTTGCACACAGC	1454						
DB	1686	GGATACCCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTTCGTTTACTTGCACACAGC	1745						
QY	1455	TGTGATGTAGTGAACCACTGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTAT	1514						
DB	1746	TGTGATGTAGTGAACCACTGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTAT	1805						
QY	1515	GAGCCATGGCGCCAAATTCCTTTCAGATTTTCATGGCTTATGACATTCGCTGTGACCG	1574						
DB	1806	GAGCCATGGCGCCAAATTCCTTTCAGATTTTCATGGCTTATGACATTCGCTGTGACCG	1865						
QY	1575	CTGTCAGTGTCTCCAGACCTGATCTCCACATGATGTCACACAGATTTTGGCCATTCAGC	1634						
DB	1866	CTGTCAGTGTCTCCAGACCTGATCTCCACATGATGTCACACAGATTTTGGCCATTCAGC	1925						
QY	1635	AGTCTTTACACAAACAGCACCATGATGTTTGGTGGTTCATTAAGTGTCTCTCTCAGCGA	1694						
DB	1926	AGTCTTTACACAAACAGCACCATGATGTTTGGTGGTTCATTAAGTGTCTCTCTCAGCGA	1985						
QY	1695	CATCTGGTATTTCACTTCGGAACAGTGTGATGCGCATCGGAGTGAACCGCTGTTTAGC	1754						
DB	1986	CATCTGGTATTTCACTTCGGAACAGTGTGATGCGCATCGGAGTGAACCGCTGTTTAGC	2045						
QY	1755	AGCAGGACCTGGTATTCGGTGTGTGTGGAACACACAGGGTTCCTCTCAGTGTATCTCTGGGC	1814						
DB	2046	AGCAGGACCTGGTATTCGGTGTGTGTGGAACACACAGGGTTCCTCTCAGTGTATCTCTGGGC	2105						
QY	1815	GCTGGCACTGATGAAACAGAGAAAAGTTAAATCAGATGTTTTCGAAAGAACTCT	1874						
DB	2106	GTTGGCACTGATGAAACAGAGAAAAGTTAAATCAGATGTTTTCGAAAGAACTCT	2165						



Qy	1875	TGACCAATGACAGATGTGACCAAGCACACAGAAATGTTACAGCTGTCACAGCCAAACACCAATGA	19334
Db	2166	TGACCAATGACAGATGTGACCAAGCACACAGAAATGTTACAGCTGTCACAGCCAAATACCAATGA	22225
Qy	1935	CTGCCACTGTCGAATGACCAATGTGTCGCCACAGAAACCAAGCTGCTCAGAAAGGCCAGAT	1994
Db	2226	CTGCCACTGTCGAATGATCACTGTGTCTCTGTGAACCAACAGTGCACAGAGGCCAGAT	2285
Qy	1995	CTCCATTTTGGTATGAGAAATTTGCCCAAGGATAACCCCTATGTACTCTGTAAACAGAA	2054
Db	2286	CTCCATTTGCCAAGTATGACAAATTTGCCCAAGGATAACCCCATGTACTCTGTCAATTAAGAA	2345
Qy	2055	GACCAGCTGCAGAGAGCTGCCCTTGGACACAGAACTGCCAGTGGAGAGCCCGGAATCAGGA	2114
Db	2346	AACAGCTGCAGAGAGCTGTGCCCTAGATCAGAAATTTGCCAGTGGAACTCGAAATCAAGA	2405
Qy	2115	GTGCATTTGCCCTGCCCGAAAATATCTGTGCAATTTGGCTGGCATTTTGGTTGGAAACTCATG	2174
Db	2406	GTGCATCGCCTTGCCCGAAAACATCTGTGGCATTTGGCTGGCAATTTGGTCGGAAACTCGTG	2465
Qy	2175	TTTGAATAATTTACTACTGCCAAGAGAAATTTATGCAATGCTAAATTTGTTCTGTAGGAACCA	2234
Db	2466	TCTGAAATCACTACTGCTTAAGGAAAATTTATGCAATGCTAAATTTGCTCTGTAGAAACCA	2525
Qy	2235	CAATGCCCTTTTGGCTTCTCTTACAAACCCAGAAAGGTAGAAATTTGTCTCTTAAGCAGCT	2294
Db	2526	CAATGCCCTTTTGGCTTCGCTTAGTCCACAGAAAGGTGAGTTTGTCTCTTAAGCAGCT	2585
Qy	2295	GCGAATAATGCAATCATCTCAGAGCATGTCCAGAGCTCACCTTTAACCCCATGGGTGGCGCT	2354
Db	2586	TCGATTAATGCAATCATCTCAGAGCATGTCCAGAGCTCACCTCTGACTCCATGGGTGGTCT	2645
Qy	2355	TCGGAAGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCATTTACAAATAGTTT	2414
Db	2646	TCGGAAGATCAATGTATCTTACTGTGCTGGAGGATATGTCTCCATTCACAAATAGTTT	2705
Qy	2415	ACTACAGTGAATGCCGTCTGAGCCAGTATGTCTGGAAATCTCTGGAAATTTTATCAGAAC	2474
Db	2706	GCTGCAGTGAATGCATCTGAGCCAGTATGTCGGGTCTCTGTGGGATCTTGTTCAGAGCC	2765
Qy	2475	CAGTACTCGGGAGCTGAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTCTGTGTA	2534
Db	2766	TAGTACTCGGGAGCTGAAGGCTGCAACTGCATCAACCCACTCAATGGAAAGTGTGTGTA	2825
Qy	2535	AAGCCCTGCAAAACCAAGTGTCTAGCAGTGCCTGAGACCACTATGTGCCCTTGAGGACAGCATG	2594
Db	2826	AAGCCCTGCAAAACCAAGTGCCTAGCAGTGCCTGAGACCACTATGTGCCCTTGAGGACAGCATG	2885
Qy	2595	TGGAGATTCACAGCGGCAGCTCTGAGTGCATGTGGTGCAGCAGCAATGAAGCAGTGTGT	2654
Db	2886	TGGTGAATGTCACAGCAGCAGCTCCGAATGCATGTGTGCAGCAACATGAAGCAGTGCCT	2945
Qy	2655	GGATCTCAATGCCATGTGGCCCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT	2714
Db	2946	AGACTCCAAATGCCTACGTAGCCTCTTCCCTTTCGGCCAGTGTATGAGTGGTATACAAT	3005
Qy	2715	GAGCACTGCCCCCTGAAATTTGTTCAAGGCTACTGTACCTGTAGTCAATTTGCTTGGAGCA	2774
Db	3006	GAGCAGCTGCCACCTGAAATTTGCTCTGGCTACTGTACCTGTGAGCCATTTGCTTGGAGCA	3065
Qy	2775	ACCAGGCTGTGGCTGTACTGTATCCAGCAATCTGCGCAACAGGAAATGCAATAGAGGG	2834
Db	3066	GCCAGGCTGTGGTGTGTACTGTATCTAGCAATCTTGGGAAAGGCAAAATGATTAGGGG	3125
Qy	2835	TTCTATTAAGGAACAGTGAAGATGCCCTTTCGCAAGCCCTCTACAGGAAATTTCTATCCACA	2894
Db	3126	AAGCTATAAAGGACCGGTGAAGATGCCCTTTCACATGCCCTCTACAGGAAATGTATCCACCA	3185
Qy	2895	GCCCTGCTCAATTTCCAGCATGTCTCTAGAGGACAGCAGATACAACTGGTCTTTTCAATCA	2954
Db	3186	GCCCTCTTGAATCCAGCATGTCTCTAGAGGACAGCAGATACAACTGGTCTTTTCAATCA	3245
Qy	2955	CTGTCCAGCTTTGCCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAA	3014

Db	3246	CTGTCCAGCTTGGCAGTGGCAATGGACACAGCAAGTGCATCAACAGAGCAATCTGTGAA	3305
Qy	3015	GTGTGAAAGCTGACCAAGCAAGCACTCGAGAACTGCAATCTCGCTTCTACGGTGA	3074
Db	3306	GTGTGAGCACTGACCAAGCAAGCACTCGAGAACTGCAATCTCGCTTCTATGGTGA	3365
Qy	3075	TCCACCAATGGAGGAAATGTCAGCCATCAAGTGCATGAGGCGACCGTCTCTGTGCAA	3134
Db	3366	CCGACTTAATGGAGGCAAAATGTGAGCCATGCAAGTGCACGGGCAACGATCTGTGAA	3425
Qy	3135	CACCAACACGGGCAAGTGTCTTGCAACCAAGAGGGCGTCAAGGGGAGCAGATGCCAGCT	3194
Db	3426	CACCAACACTGGCAAGTGCTTCTGTACCAACAGGGGCGTCAAGGGGAGGAGTGCCAGCT	3485
Qy	3195	ATGTGAGGTAGAAAATCGATACCAAGAAACCCCTCTCAGAGGAACATGTTATTATCTCT	3254
Db	3486	ATGTGAGGTAGAAAATCGATACCAAGGGAACCCCTCTCAAGGAACATGTTATTATCCCT	3545
Qy	3255	TCTTATTGACTATCAGTTCACTTTAGTCTATGCCAGGAAGATGATCGTATTATCACAGC	3314
Db	3546	TCTCATTTCACTATCAGTTTCACTTTTAGCCCTGTCCAGGAAGACACCGCTATTATCACAGC	3605
Qy	3315	TATCAATTTTGTGCTACTCTCTGACGAACAAACACAGGATTTTGACATGTTCAATATGC	3374
Db	3606	CATCAATTTTGTGCTACTCTCTGATGAACAAACACAGGATTTTGACATGTTCAATATGC	3665
Qy	3375	CTCCAAGAAATTTCAACCTCAACATCACTCGGCGTCCAGTTTCTCAGCTGGAACCCAGGC	3434
Db	3666	CTCCAAAACCTTCAACCTCAACATCACTCGGCGTCCAGTTTCCAGCTGGAACCCAGAC	3725
Qy	3435	TGGAGAAGAGATGCTGCTGTGTTTCAAAAACCAACATTAAGAGATACAAAGATAGTTCTC	3494
Db	3726	TGGAGAAGAGTGTGCTGTGTTTCAAAAACCAACATCAAGGAATACAAAGATAGTTCTC	3785
Qy	3495	TAATGAGAAGTTGAATTTGCCAACCCCAATATACATTTCTTTGTTATGTAGTAA	3554
Db	3786	TAATGAGAAATTTGAATTTGCCAACCCCAATATACATTTCTTTGTTATGTAGTAA	3845
Qy	3555	TTTCACCTGGCCCATCAAAATTC	3578
Db	3846	TTTCACCTGGCCCATCAAAATTC	3869
RESULT 12			
AB093245		5683 bp	linear
LOCUS	AB093245		mRNA
DEFINITION	Mus musculus mRNA for mKIAA0548 protein.		
ACCESSION	AB093245		
VERSION	AB093245.1	GI:26006172	
KEYWORDS	FLI_CDNA.		
SOURCE	Mus musculus		
ORGANISM	Mus musculus (house mouse)		
REFERENCE			
AUTHORS	Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Hara, Y., Nagase, T., Ohara, O. and Koga, H.		
TITLE	Prediction of the coding sequences of mouse homologues of KIAA gene: I. The complete nucleotide sequences of 100 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 5683)		
AUTHORS	Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.		
TITLE	Direct Submission		
JOURNAL	Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918)		
COMMENT	The CREATE program supported by Japan science and technology		

15 "\*\*\*\*\*"  
The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'-&

COMMENT

3'-end one pass sequencing.	
FEATURES	Location/Qualifiers
source	1..5683
	/organism="Mus musculus"
	/mol_type="mRNA"
	/db_xref="taxon:10090"
	/clone="mbg03041"
	/tissue_type="brain"
	/dev_stage="adult"
	/note="vector:modified pBC SK+."
gene	1..5683
	/gene="mkiAA0548"
CDS	<1..3181
	/gene="mkiAA0548"
	/note="CDS is predicted by in silico analysis. Start codon is not identified."
	/codon_start=2
	/evidence="not experimental"
	/product="mkiAA0548 protein"
	/protein_id="BAC41429.1"
	/db_xref="GI:26006173"
	/translation="ADLTSLREWLPLNHSVNSVVRVYGHSLALHKDKIYNYGKIDST GNTNELRVPHIHNESWLLTPRAKQOYAVVGHSAHIVTLASGRVVMVIFGHCPYIG YISVQEDYDEKNTSILHTQALVQGGVSHSVYDDRTKALYVHGVIYKAFSGKRL ADLYDYDVTQWTLTKDSRFYRHLTAIVISGTMVFGNTHNDLSMHSKAFSS DFMAYDIACDRWVLPRLPHLDVNRFGHSAVLVNSTMYVFGFNFSLSDVLVPTLE QDAHRSEACVAAGPIRCLWDTQSRTSWELEAOEKLKSECFKRLDHDRC DDHTCYSCNTANDCHMNDHCVPVNSHCTEGQISIAKVESCPKDNPMVYCNKKTSC RSCALQNCWEPRENEICIALPENICNGHVLVNSCLIKITAKENYDNKLUSCRNHN AFILASTSQKVFQKRLQMOSSMSKLTLPWVGRKINVSITWCMEDMSPTNS LLQMFSEPSDAGFCILBEPSTRGLKAATCINPLNSVCERPANHSAKQRTPCALR TAGCECTSSSECCWNNKQCVDSNAYVSPFGQCEWYTWSSCPPECCSGYCTCS HCLQPCGNCCTDPSNTGKCKIEGSKYKVPMSOASAGNVYPOPILNNSMCLEDSR YNSFPCHPACQNGHSCKINOSVCEKEDLTGKHCECTCISFVGDPTNGKQCPCK CNGHSLCNTNGKCTTKGVKAGDECOLCEVENRYOQNPLKGTCTYTLILDYQTFES LSQEDRYTAINFPAIPBQNRDLDMFINASKNFNLNTWATISFPAGTQTGEVPV SKTTIKEYKDSFNSKDFPNHNTIFFVYNSNFTWPIKIQIAFSOHSNFMOLVQFFV TFSCFLSLLAAVVMKIQSCWASRRBQLREMOQWASFPFASVNVALREDEPP DLIGSGIKITVPKLEPCFGKNAVLVSFVRLPRGLGGIPPPGQSLAVASALVDIS QQMPIVYKEKSGAVNRKQPPAQPGTCI"
ORIGIN	
Query Match. 56.2%; Score 2694.6; DB 10; Length 5683;	
Best local Similarity 90.4%; Pred. No. 0;	
Matches 2877; Conservative 0; Mismatches 304; Indels 0; Gaps 0;	
Qy	888 AGCTATGACCTGCTCTTAGGAGTGGCTTCCACTAAACCGTTCTGTGAAACAATGTGGT 947
Db	1 AGCGTATGACCTGACCTTAGGGAATGGCTTCCACTAAACCACTTCTGTGAAACAGTGTGT 60
Qy	948 TGTATGATATGTCATCTTTGGCATATATCAAGAGATAAATAATTAATGTATGGAGGAAA 1007
Db	61 TGTAAAGATATAGTCATCTTTGGCATATATCAAGAGATAAATAATTAATGTATGGAGGAAA 120
Qy	1008 AATTGATTCAACTGGGAATGTGACCAATGAGTTGAGAGTTTTTTCACATTCATAATGAGTC 1067
Db	121 AATTGATTCAACAGGAACTGACCAATGAGCTGAGAGTATTTTCATATTCATAATGAAATC 180
Qy	1068 ATGGGTGTTGTTGACCCCTTAAGCAAGGAGCAGTATGACAGTGGTTGGSCACTCTGCA 1127
Db	181 ATGGGTATTTGTTAACTCCGAAAGCTAAGGATCAGTATGACAGTGGTTGGCACTCAGCACA 240
Qy	1128 CATGTGTACACTAAGAAAGCGGAGTGGTCATGCTGGTCACTCTTTGGTCACTGCCCTCT 1187
Db	241 CATGTGTACACTGACATCTGGCGTGGTGGTCATGTTGGTCACTCTCGGTCATGGCCACT 300
Qy	1188 CTATGGATATATAAGCAATGTGAGGAATATGATTTGGATAAGAAACAATGAGGATATAT 1247
Db	301 CTATGGATATATAAGCGTTGTGAGGAATATGATTTGGATAAGAAACAATGAGGATATAT 360
Qy	1248 ACACACCCAGGTGCCCTTGTGCAAGGGGTACGGCCATGACAGTGGTTTACGACCATAG 1307
Db	361 ACATACCTCAGGGTGCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTATTATGATGACAG 420
Qy	1308 GACCAAGGCGCCCTATACGTTTCATGTTGCTGCTACAGGCTTTTCTAGTCCCAATAGTACCGGCT 1367
Db	421 GACCAAGGCTCTGTACGTTTCATGTTGCTGCTACAGGCTTTTCTAGTCCCAATAGTACCGGCT 480
Qy	1368 TGCAGATGATCTCTACCGCATATGATGTGGATACCCAGATGTGGACCACTTCTTAAGGACAG 1427
Db	481 TGCAGATGACCTCTACAGATACGATGTGGATACCTCAGATGTGGACCACTTCTTAAGGACAG 540
Qy	1428 CCGATTTTTCCGTTACTTTGTCACACAGCTGTGATAGTGGAGGAAACCAATGCTGGTGTGG 1487
Db	541 CCGATTTTTCCGTTACTTTGTCATACAGCTGTGATAGTGGAGGAAACCAATGCTGGTGTGG 600
Qy	1488 GGGAAACACACACATGACATCTATGAGCCATGGCGCCCAATGCTTCTCTTCAGATTT 1547
Db	601 AGGAAACACACACATGACATCTTCCATGAGCCAGGTGCCCAATGCTTCTCTTCGACCTT 660
Qy	1548 CATGSCCTTATGACATGCTGTGACCGCTGTGACCGTGTGCTGTCTTCCAGACCTGATCTCCACCA 1607
Db	661 CATGSCCTTATGACATGCTGTGACCGATGCTGTGACCGTGTGCTGTCTTCCAGACCTGATCTCCATCA 720
Qy	1608 TGATGTCAACAGATTTGGCCATTCAGCAGCTTTACACAAACAGCACCATGTATGTGTTCCG 1667
Db	721 TGATGTCAACAGATTTGGCCATTCAGCAGCTTTACACAAACAGCACCATGTATGTGTTCCG 780
Qy	1668 TGGTTTCAATAGTCTCTCTCTCAGCAGATCCTGTGTATTTACCTCGGAAACAGTGTGATGC 1727
Db	781 CGGCTTCAACAGCTCTCTCTCAGTGTGCTGTCTTTACCTTGGAGCAGTGTGATGC 840
Qy	1728 GCATCGGAGTGAAGCCGCTTGTGATGACAGGACCTTGTGATTCGGTGTGTTGGAAACAC 1787
Db	841 ACACCGCAGTCAAGCTGCTGTGTGGCAGCAGGACCTGATTCGGTGTCTGTGGGACAC 900
Qy	1788 AGGCTGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAAAGTTAAA 1847
Db	901 ACAGTGTCTCAGTGTATCTCTCGGAGTTGGCAACTGAGAAACAAGCAAGAAAGTTAAA 960
Qy	1848 ATCAGAAATTTTTTCCAAAAGAACTTTGACCAATGACAGATGTGACAGACACAGATTG 1907
Db	961 ATCAGAGTGTTTTTTCTAAAAGAACCTTTGACCAATGACAGATGTGACCAAGACACAGATTG 1020
Qy	1908 TTACAGCTGCACAGCCCAACCAATGACTGCACCTGTGCAAGTGAACATTTGTGCCACAG 1967
Db	1021 TTACAGCTGCACAGCCCAATGACTGCACCTGTGCAAGTGAACATTTGTGCCCTGT 1080
Qy	1968 GAACCAAGCTGTCTCAGAAAGCCAGATCTCCATTTTATAGTATGAGAAATTTGCCCAAGGA 2027
Db	1081 GAACCAAGCTGTGCACAGAAAGCCAGATCTCCATTTGCCAAGTATGAGAGTTGCCCAAGGA 1140
Qy	2028 TAAACCTTATGTACTACTGTAAACAGAACCAAGCTGCAGGAGCTGCTCCCTGGACACGAA 2087
Db	1141 TAAACCTTATGTACTACTGTAAACAGAACCAAGCTGCAGGAGCTGCTCCCTAGACACGAA 1200
Qy	2088 CTGCCAGTGGAGAGCCCGGAATCAGGAGTGCATTTGCCCTCGCCGAAATATCTGTGSCAT 2147
Db	1201 CTGCCAGTGGAGAGCCCGGAATCAAGAGTGCATCGCCCTCGCCGAAATATCTGTGSCAA 1260
Qy	2148 TGGCTGGCAATTTGGTTGAAACTCATGTTTGAATAATTTACTACTGCCAAGAGAAATTA 2207
Db	1261 TGGCTGGCAATTTGGTTGAAACTCATGTTTGAATAATTTACTACTGCCAAGAGAAATTA 1320
Qy	2208 CAATGCTAAATTTGTTCTGTAGGAACCAATGCTCCCTTTTGGCTTCTCTTACAAACCCAGAA 2267
Db	1321 CAATGCTAAATTTGTTCTGTAGGAACCAATGCTCCCTTTTGGCTTCTCTTACAAACCCAGAA 1380
Qy	2268 GAAGGTGAAATTTGCTTAAAGAGCTGCGAAATATGCAATGATCATCTCAGAGCATGTCCAA 2327
Db	1381 GAAGGTGAGATTTGCTTAAAGAGCTTCAATTAATGCAATCATCTCAAAAGTATGTCCAA 1440
Qy	2328 GCTCACCTTAAACCCCATGGGTGGCCCTTCGGAAGATCAATGTGCTCTTACTGTGCTGGGA 2387
Db	1441 GCTCACCTTAAACCCCATGGGTGGCTTTCGGAAGATCAATGTGCTCTTACTGTGCTGGGA 1500

QY 2388 AGATATGTCCCAATTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCCAAGTATGC 2447  
Db 1501 GGATATGTCTCCCAATTTACAAATAGTTTGTGCGAGTGGATGCCATCTGAGCCCAAGTATGC 1560  
QY 2448 TGGATTTCTGGGAATTTTATCAGAAACCAAGTACTCGGGGACTGAAAGGCTGCAACCTGCGAT 2507  
Db 1561 TGGCTTCTGGGATCTTTGTACAGAGCTAGTACTCGGGGATTAAGGCTGCAACCTGCGAT 1620  
QY 2508 CAACCACTCAATGGTAGTCTGTGAAGGCTGCAACCAAGTACTGAGAGTGGCG 2567  
Db 1621 CAACCCCTCTCAATGGCAGCGCTGTGAAGGCTGCAACCAAGTACTGAGAGTGGCG 1680  
QY 2568 GACACCATGTGCTTGGAGCAGCATGTGAGAGTGCACAGCGGCGAGCTCTGAGTGCAT 2627  
Db 1681 GACACCATGTGCTTGGAGCAGCATGTGAGAGTGCACAGCGGCGAGCTCTGAGTGCAT 1740  
QY 2628 GTGGTCAGCAACATGAAGCAGTGTGTGGAATCCAAATGCTATGTGGCTCTCTCCCTTT 2687  
Db 1741 GTGGTCAGCAACATGAAGCAGTGTGTGGAATCCAAATGCTATGTGGCTCTCTCCCTTT 1800  
QY 2688 TGGCCAGTGTATGGATGGTATACGATGAGCAGCTGCCCTGCAAAATTTGTTCAAGCTA 2747  
Db 1801 TGGCCAGTGTATGGATGGTATACGATGAGCAGCTGCCCTGCAAAATTTGTTCAAGCTA 1860  
QY 2748 CTGTACCTGTAGTCAATGTGTGAGCAACAGGCTGTGGGTGTGTACTGATCCCAAGCAA 2807  
Db 1861 TTGTACCTGTAGTCAATGTGTGAGCAACAGGCTGTGGGTGTGTACTGATCCCAAGCAA 1920  
QY 2808 TACTGGCAAGGGAATTCATAGAGGTTCTATTAAGACCAAGTGAAGTGCCTTCGCA 2867  
Db 1921 TACTGGCAAGGGAATTCATAGAGGTTCTATTAAGACCAAGTGAAGTGCCTTCGCA 1980  
QY 2868 AGCCCTTACAGGAAATTTCTATCCAGCCCTGCTCAATTCAGCATGTGTAGAGGA 2927  
Db 1981 GGCTCTGACAGGAATGTGTATCCAGCCCTCTGAACTCCAGCATGTGTAGAGGA 2040  
QY 2928 CAGCAGATCAACTGTGTTTCAATCTACTGTCCAGTGTGCAATGCAAGCCCAAGTAA 2987  
Db 2041 CAGCAGATCAACTGTGTTTCAATCTACTGTCCAGTGTGCAATGCAAGCCCAAGTAA 2100  
QY 2988 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGACCTGACCAAGCAAGCACTGCGA 3047  
Db 2101 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGACCTGACCAAGCAAGCACTGCGA 2160  
QY 3048 GACCTGATATCTGGCTTCTACGGTGTATCCCAATGGAGGAAATGTGAGCCATGCAA 3107  
Db 2161 GACCTGATATCTGGCTTCTACGGTGTATCCCAATGGAGGAAATGTGAGCCATGCAA 2220  
QY 3108 GTGCAATGGGACCGCTCTGTGCAACCAACCAAGGCAAGTGTCTGTGCAACCAAC 3167  
Db 2221 GTGCAATGGGACCGCTCTGTGCAACCAACCAAGGCAAGTGTCTGTGCAACCAAC 2280  
QY 3168 GGGGCTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAATCGATACCAAGGAACCC 3227  
Db 2281 AGGTGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAATCGATACCAAGGAATCC 2340  
QY 3228 TCTCAGAGGAACATGTTATATATCTCTTATGACTATCAGTTCACCTTTAGTCTATC 3287  
Db 2341 TCTCAGAGGAACATGTTATATATCTCTTATGACTATCAGTTCACCTTTAGTCTATC 2400  
QY 3288 CCAGAGATGATCGCTATTACAGCTATCAATTTTGGCTACTCTGACCAACAA 3347  
Db 2401 CCAGAGATGATCGCTATTACAGCTATCAATTTTGGCTACTCTGACCAACAA 2460  
QY 3348 CAGGGATTTGACATGTTTCATATGCTTCAAGAAATTTCAACCTCAACATCACTGGG 3407  
Db 2461 CAGGGATTTGACATGTTTCATATGCTTCAAGAAATTTCAACCTCAACATCACTGGG 2520  
QY 3408 TGCAGGTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCAA 3467  
Db 2521 CACCAGCTTCCAGCCGGAACCCAGCTGGAGAGAGGTCCTGTGTTTCAAAACCAA 2580  
QY 3468 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTTTTCGCAACCAACCCAA 3527

Db 2581 CATCAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTTCGACCATCCAA 2640  
QY 3528 TATCATTCTTTTGTATGTCAATAATTTCACTGGCCCATCAAAATTCAGATTGCCTT 3587  
Db 2641 CATCATTCTTTTGTATGTCAATAATTTCACTGGCCCATCAAAATTCAGATTGCCTT 2700  
QY 3588 CTCTCAGCACAGCAATTTTATGGACCTGGTACAGTCTTCTGTGACTTCTTCAAGTTGTT 3647  
Db 2701 CTCCAGCACAGCAATTTTATGGACCTGGTACAGTCTTCTGTGACTTCTTCAAGTTGTT 2760  
QY 3648 CCTCTCTTTTGTCTGCTGTGTGGTTTGAAGATCAAAAGTTTGTGGGCTC 3707  
Db 2761 CCTCTGCTGTCTTCTGTGTGGTCTGAGTGGTCTGCAAGATCAAGCAGAGCTTGGGCATC 2820  
QY 3708 CAGAGTATAGAGCAACTTCTTCGAGAGATGCAACAGATGCGCAGCGCTCTTGGCTC 3767  
Db 2821 CAGGCGGAGAGCAACTTCTTCGAGAGATGCAACAGATGCGCAGCGCTCTTGGCTC 2880  
QY 3768 TGTAAATGTCCTTGGAAACAGATGAGAGCTCTCTGATCTTATTTGGGGGAGTATAAA 3827  
Db 2881 TGTAAACGTTGCTTGGAAACAGATGAGAGCTCTCTGATCTCATTTGGGGGAGTATAAA 2940  
QY 3828 GACTGTTCCCAACCCATTCACATGACCTGGAGCCGTGTTTGGCAACAAAGCCGTGCTCTC 3887  
Db 2941 GACCCGTTCTTAAGCCCATTTGCCCTGGAGCCCTGCTTTGGTAAACAAAGCCGTCTCTC 3000  
QY 3888 TGCTGTTCTGAGGCTCCCTCGAGGCTCGGTGGATCCCTCTCTGGGAGTCAAGGTCT 3947  
Db 3001 TGTTCTGAGGCTCCCTCGAGGCTCGGTGGAGTCAAGGTCTCTCTCTGAGTCAAGGTCT 3060  
QY 3948 TGCTGTCGACGCGCTTGGTGGACATTTCTCAGCAGATGCCGATAGTGTCAAGAGAA 4007  
Db 3061 CGCTGTGCGCAGTGGCTTGGTGGACATTTCTCAGCAGATGCCGATAGTGTCAAGAGAA 3120  
QY 4008 GTCAGAGCCCTGAGAAACCGAGCAGCAGCCCTGTCACAGCTGGGACCTGCATCTG 4067  
Db 3121 GTCAGAGCTGTAGAAACCGAGCAGCAGCCCTGTCACAGCTGGGACCTGCATTTG 3180  
QY 4068 A 4068  
Db 3181 A 3181

RESULT 13  
AR164818  
LOCUS  
DEFINITION Sequence 18 from patent US 6274339.  
ACCESSION AR164818  
VERSION AR164818.1 GI:15238029  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
source  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Query Match 50.4%; Score 2051.8; DB 6; Length 2625;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 2089; Conservative 0; Mismatches 62; Indels 0; Gaps 0;  
QY 15 GCGCGCACTGAGGCAAGCTCAGAGGAGGAGCAGCGCGGCGGAGCGCTCGCGGGCAG 74  
Db 234 GGAGGCCGAGGCGCTGCGGTGGCGGCGGCGGTGTCCGCTCGCGCGAGCGGAGCCAA 293



```

VERSION AR164815.1 GI:16238024
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 6370)
AUTHORS Moore,K. and Ragle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
JOURNAL weight disorders, including obesity
PATENT Patent: US 6274339-A 12 14-AUG-2001;
FEATURES Location/Qualifiers
source 1..6370
/organism="unknown"
/mol_type="unassigned DNA"
ORIGIN
Query Match 45.6%; Score 1853; DB 6; Length 6370;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;
QY 2035 ATGTACTACTGTAAAGAACACAGCTGAGAGAGCTGTGCCCTGGACCAAGTCCAG 2094
DB 1 ATGTACTACTGTAAAGAACACAGCTGAGAGAGCTGTGCCCTGGACCAAGTCCAG 60
QY 2095 TGGAGCGCCGGAAATCAGAGATGCATTGGCCCTGCCGAAATATCTGTGCCATTTGGCTGG 2154
DB 61 TGGAGCGCCGGAAATCAGAGATGCATTGGCCCTGCCGAAATATCTGTGCCATTTGGCTGG 120
QY 2155 CATTGGTTGGAAATCATGTTTGAATAATTAATCTACTGCGAAGAGAAATTAAGCAATGCT 2214
DB 121 CATTGGTTGGAAATCATGTTTGAATAATTAATCTACTGCGAAGAGAAATTAAGCAATGCT 180
QY 2215 AATTTGTTCTGTAGGAACCAATGCCCTTTGGCTTCTTTACAAACCCAGAAAGGTA 2274
DB 181 AATTTGTTCTGTAGGAACCAATGCCCTTTGGCTTCTTTACAAACCCAGAAAGGTA 240
QY 2275 GAATTTGTTCTTAAGCAGCTGCGAATAATGCAATCATCTCAGAGCATGTCCAAAGTCAAC 2334
DB 241 GAATTTGTTCTTAAGCAGCTGCGAATAATGCAATCATCTCAGAGCATGTCCAAAGTCAAC 300
QY 2335 TTAACCCCATGGTGGCCCTT---CGGAAGATCAATGTGTCTTACTGG---TGCTGGG 2386
DB 301 TTAACCCCATGGTGGCCCTTCCGGAAGGTCAATGTGTCTTACTGCTGCTGCTGGGGAAG 360
QY 2387 AAGATATGTCCTCCATTTACAAATAG-TTTTACTACAGTGGATGCC---GTCTGAGCCAC 2441
DB 361 GATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCCCGTCTTGAGGCCAC 420
QY 2442 TGATGCT--GGATCTGTGGAATTTT-----ATCAGAACCCAGTACTCGGGGACTGAAG 2494
DB 421 TGTGTCTTGGATTTCTGTGGGAATTTTATTTCAGGAACCCAGTTTACTTCGGGGACTGAAG 480
QY 2495 CTGCAACCTGCAT--CAACCCACTCAATGTTAGTGTGTGTAAGGCCCTGCAACACACAG 2552
DB 481 CTGCAACCTGCATTTCAACCCACTYMAATGTTAGTGTGTGTAAGGCCCTGCAACACACAG 540
QY 2553 TGCTAA--GCAGTGGCGACACCATGTGCTTTGAGGACAGCATGTGAGATTTGACCCAGCG 2611
DB 541 TGCTAAGCGAGTGGCGACACCATGTGCTTTGAGGACAGCATGTGAGATTTGACCCAGCG 600
QY 2612 GCAGCTGTGATGTGTGTGCGACCAATGAGCATGTGTGGAATTCGATCCAAATGCTCATG 2671
DB 601 GCAGCTGTGATGTGTGTGCGACCAATGAGCATGTGTGGAATTCGATCCAAATGCTCATG 660
QY 2672 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 2731
DB 661 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCCCCTG 720
QY 2732 AAAATTTGTCAGGCTACTGTACCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGT 2791
DB 721 AAAATTTGTCAGGCTACTGTACCTGTAGTCAATTTGCTTGGAGCAACAGGCTGTGGCTGT 780
QY 2792 GTACTGTATCCAGCAATACTTGGCAAGGGAAATGTCATAGAGGGTTCTATAAAGGACCAG 2851
781 GTACTGTATCCAGCAATACTTGGCAAGGGAAATGTCATAGAGGGTTCTATAAAGGACCAG 840
2852 TGAAGATGCCCTTTCGAAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCA 2911
841 TGAAGATGCCCTTTCGAAAGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCA 900
2912 GCATGTCTGTAGAGGACAGAGATACAACTGGTCTTTTCACTCACTGTCAGCTTGGCCAAT 2971
901 GCATGTCTGTAGAGGACAGAGATACAACTGGTCTTTTCACTCACTGTCAGCTTGGCCAAT 960
2972 GCAACGGCCACAGTAAATGTCATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACCA 3031
961 GCAACGGCCACAGTAAATGTCATCAATCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACCA 1020
3032 CAGGCAAGCATCTGGAGACCTGCATATCTGGCTTCTAGGCTGATCCCAACCAATGGAGGA 3091
1021 CAGGCAAGCATCTGGAGACCTGCATATCTGGCTTCTAGGCTGATCCCAACCAATGGAGGA 1080
3092 AATGTTCAGCCATGCAAGTGGCAACGCGTCTCTGTGCAACCAACCAACGCGGCAAGT 3151
1081 AATGTTCAGCCATGCAAGTGGCAACGCGTCTCTGTGCAACCAACCAACGCGGCAAGT 1140
3152 GCTTCTGCACCAACCAAGGGCGTCAAGGGGGACGAGTGCAGCTATGTGAGTAGAAATC 3211
1141 GCTTCTGCACCAACCAAGGGCGTCAAGGGGGACGAGTGCAGCTATGTGAGTAGAAATC 1200
3212 GATACCAAGGAACCCCTCTCAGAGGAACATGTTATATATCTCTTCTTATTGACATATCAGT 3271
1201 GATACCAAGGAACCCCTCTCAGAGGAACATGTTATATATCTCTTCTTATTGACATATCAGT 1260
3272 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTAACAGCTATCAATTTTGTGGCTA 3331
1261 TCACCTTTAGTCTATCCAGGAAGATGATCGTATTAACAGCTATCAATTTTGTGGCTA 1320
3332 CTCTCGACGAACCAACAGGAGTTTGGACATGTTTCATCAATGCGCTCCAAAGATTTCAACC 3391
1321 CTCTCGACGAACCAACAGGAGTTTGGACATGTTTCATCAATGCGCTCCAAAGATTTCAACC 1380
3392 TCAACATCACCTGGGCTGCCAGTCTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTG 3451
1381 TCACATCACCTGGGCTGCCAGTCTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTG 1440
3452 TTGTTTCAAAAACCAACATTAAGAGATCAAGATAGTTTCTCTAATGAGAGTTTGAAT 3511
1441 TTGTTTCAAAAACCAACATTAAGAGATCAAGATAGTTTCTCTAATGAGAGTTTGAAT 1500
3512 TTGCAACCAACCCAAATATCACTTTCTTTGTTTATGTCAGTAAATTCACCTGGCCATCA 3571
1501 TTGCAACCAACCCAAATATCACTTTCTTTGTTTATGTCAGTAAATTCACCTGGCCATCA 1560
3572 AAAATTGAGATTGCCCTTCTCAGCAACAGCAATTTTATGGAAGTGTGATGAGTGGCCA 3631
1561 AAAATTGAGATTGCCCTTCTCAGCAACAGCAATTTTATGGAAGTGTGATGAGTGGCCA 1620
3632 CTTTCTTCTCAGTTGTTTCTCTTTGCTTCTGCTGGCTGCTGCTGTTTGGAGATCAAAC 3691
1621 CTTTCTTCTCAGTTGTTTCTCTTTGCTTCTGCTGGCTGCTGCTGTTTGGAGATCAAAC 1680
3692 AAAATTGTTGGGCTTCCAGACCTGAGAGCAACTTTCTTCGAGAGATGCAACAGATGGCCA 3751
1681 AAAATTGTTGGGCTTCCAGACCTGAGAGCAACTTTCTTCGAGAGATGCAACAGATGGCCA 1740
3752 GCGCTCCCTTTGCTCTGTAAATGTCGCTTGGAAACAGATGAGAGCTCTCTGATCTTA 3811
1741 GCGCTCCCTTTGCTCTGTAAATGTCGCTTGGAAACAGATGAGAGCTCTCTGATCTTA 1800
3812 TTGGGGGAGATTAAGACTGTTCCAAACCCATTCGACTGGAGCCGTGTTTGGCAACA 3871
1801 TTGGGGGAGATTAAGACTGTTCCAAACCCATTCGACTGGAGCCGTGTTTGGCAACA 1860
3872 AAGCGCTGTCTCTCTGTGTGTTGTGAGGCTCCCTCGAGGGCTGGGTGGCATCCCTCCTC 3931
```

```

Db 1861 AAGCGGTGCTCTCTGTGTGTTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTC 1920
QY 3932 CTGGGCGAGTCAGGTCCTGTCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 3991
Db 1921 CTGGGCGAGTCAGGTCCTGTCTGTGGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1980
QY 3992 TAGTGTAACAGGGAAGTCAGGAGCCGTGAGAAACCGAAGACAGCAGCCCTCTGCACAGC 4051
Db 1981 TAGTGTAACAGGGAAGTCAGGAGCCGTGAGAAACCGAAGACAGCAGCCCTCTGCACAGC 2040
QY 4052 CTGGGAGCTGCATCTGA 4068
Db 2041 CTGGGAGCTGCATCTGA 2057

RESULT 15
AK000356 2578 bp mRNA linear PRI 13-SBP-2003
DEFINITION Homo sapiens cDNA FLJ20349 fis, clone HEP13957, highly similar to
AB011120 Homo sapiens mRNA for KIAA0548 protein.
ACCESSION AK000356
VERSION AK000356.1 GI:7020387
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,
Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,
Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,
Nakamura,Y., Isogai,T. and Sugano,S.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 2578)
Sugano,S., Suzuki,Y., Ota,T., Obayashi,M., Nishi,T., Isogai,T.,
Shibahara,T., Tanaka,T. and Nakamura,Y.
Direct Submission
Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,
University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,
Minato-ku, Tokyo 108-8639, Japan (E-mail:flcdna@ims.u-tokyo.ac.jp,
Tel:81-3-5449-5286, Fax:81-3-5449-5416)
NEDO human cDNA sequencing project supported by Ministry of
International Trade and Industry of Japan; cDNA full insert
sequencing: Research Association for Biotechnology; cDNA library
construction, 5'- & 3'-end one pass sequencing: Department of
Virology and Human Genome Center, Institute of Medical Science,
University of Tokyo (partly supported by Science and Technology
Agency). Location/Qualifiers
FEATURES
source 1. .2578
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="HEP13957"
/cell_line="HepG2"
/call_type="Hepatoma"
/clone_lib="HEP"
/notes="cloning vector pME18SFL3"
misc_feature 1. .2578
/notes="highly similar to AB011120 Homo sapiens mRNA for
KIAA0548 protein"
ORIGIN
Query Match 41.7%; Score 1697; DB 9; Length 2578;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1697; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2372 CCTACTGGTCTGGGAGATATGTCCTTACAAATAGTTTACTACAGTGGATGCCGT 2431
Db 1 CCTACTGGTCTGGGAGATATGTCCTTACAAATAGTTTACTACAGTGGATGCCGT 60
QY 2432 CTGAGCCCACTGATGCTGGATTCTGTGGAAATTTATCAGAACCCAGTACTCGGGGACTGA 2491

```

```

Db 61 CTGAGCCCAAGTGATGCTGGATTCTGTGGAAATTTATCAGAAACCCAGTACTCGGGGACTGA 120
QY 2492 AGGCTGCAACCTGTCATCAACCCACTCAATGTGTGTCTGTGAAAGGCTGCAAAACACA 2551
Db 121 AGGCTGCAACCTGTCATCAACCCACTCAATGTGTGTCTGTGAAAGGCTGCAAAACACA 180
QY 2552 GTGCTAAGCAGTGGCGGACACCATGTGCTTGGAGACAGCATGTGGAGATTGCAACGCG 2611
Db 181 GTGCTAAGCAGTGGCGGACACCATGTGCTTGGAGACAGCATGTGGAGATTGCAACGCG 240
QY 2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATCAAGCAGTGTGTGGACTCCCAATGCCATG 2671
Db 241 GCAGCTCTGAGTGCATGTGGTGCAGCAACATCAAGCAGTGTGTGGACTCCCAATGCCATG 300
QY 2672 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGATGGTATACGATGACGACCTGCCCCCTG 2731
Db 301 TGGCTCTCTTCCCTTTTGGCCAGTGTATGGATGGTATACGATGACGACCTGCCCCCTG 360
QY 2732 AAAATTGTTCAAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACCAAGGCTGTGGCTGGT 2791
Db 361 AAAATTGTTCAAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACCAAGGCTGTGGCTGGT 420
QY 2792 GTACTGATCCAGCAATACTGTGGCAAGGAAATGATAGAGGTTCTTAAAGGACAG 2851
Db 421 GTACTGATCCAGCAATACTGTGGCAAGGAAATGATAGAGGTTCTTAAAGGACAG 480
QY 2852 TGAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCA 2911
Db 481 TGAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCA 540
QY 2912 GCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTCAITCACTGTCCAGCTTGCCAAT 2971
Db 541 GCATGTGTCTAGAGACAGCAGATACAACTGGTCTTTCAITCACTGTCCAGCTTGCCAAT 600
QY 2972 GCACGGCCACAGTAATGTCATCAATCAGAGCATCTGTGAGAGTGTGAGACCTGACCA 3031
Db 601 GCACGGCCACAGTAATGTCATCAATCAGAGCATCTGTGAGAGTGTGAGACCTGACCA 560
QY 3032 CAGGCAAGCACTGGGAGACCTGTCATATCTGGCTTCTACGGTGTATCCCAACCAATGGAGGA 3091
Db 661 CAGGCAAGCACTGGGAGACCTGTCATATCTGGCTTCTACGGTGTATCCCAACCAATGGAGGA 720
QY 3092 AATGTCCAGCATGCAAGTGCATGGCCAGCTCTGTGCAACCAACCAACAGGGGAAGT 3151
Db 721 AATGTCCAGCATGCAAGTGCATGGCCAGCTCTGTGCAACCAACCAACAGGGGAAGT 780
QY 3152 GCTTCTGCACCAACCAAGGGCGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAAATC 3211
Db 781 GCTTCTGCACCAACCAAGGGCGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAAATC 840
QY 3212 GATACCAAGGAAACCTCTCAGAGGAACATGTTATATCTCTTCTTATTGACTATCAGT 3271
Db 841 GATACCAAGGAAACCTCTCAGAGGAACATGTTATATCTCTTCTTATTGACTATCAGT 900
QY 3272 TCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGAGCTATCAATTTTGTGGCTA 3331
Db 901 TCACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGAGCTATCAATTTTGTGGCTA 960
QY 3332 CTCCTGAGAAACAAAACAGGATTTGGACATGTTTCATCAATGCTCCCAAGAAATTTCAACC 3391
Db 961 CTCCTGAGAAACAAAACAGGATTTGGACATGTTTCATCAATGCTCCCAAGAAATTTCAACC 1020
QY 3392 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTG 3451
Db 1021 TCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCTG 1080
QY 3452 TTGTTTCAAAAACCAACATTAAAGAGTACAAAGATGTTTCTCTTAATGAGAAGTTGATT 3511
Db 1081 TTGTTTCAAAAACCAACATTAAAGAGTACAAAGATGTTTCTCTTAATGAGAAGTTGATT 1140
QY 3512 TTGCGCAACCAACCAATATCACTTCTTCTTGTATGTCAGTAAATTTTCACTGCGCCATCA 3571

```

Db 1141 TTGCGAACACCCAAATATACACTTTCTTTGTTATGTAGTAATTTACACTGCGCCCATCA 1200  
QY 3572 AAATTGAGATTGCTTTCTCTGACACAGCAATTTTATGAGACCTGGTACAGTTCTTGGTGA 3631  
Db 1201 AAATTGAGATTGCTTTCTCTGACACAGCAATTTTATGAGACCTGGTACAGTTCTTGGTGA 1260  
QY 3632 CTTTCTTCAGTTGTTTCCCTCTCTTTGCTCTGCTGGTGGCTGCTGGTTGGAAGATCAAC 3691  
Db 1261 CTTTCTTCAGTTGTTTCCCTCTCTTTGCTCTGCTGGTGGCTGCTGGTTGGAAGATCAAC 1320  
QY 3692 AAAGTTGTTGGGCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 3751  
Db 1321 AAAGTTGTTGGGCTCCAGACGTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 1380  
QY 3752 GCGGTCCCTTTGCTCTGTAATGTGCGCTTGGAACAGATGAGGAGCCTCCTGATCTTA 3811  
Db 1381 GCGGTCCCTTTGCTCTGTAATGTGCGCTTGGAACAGATGAGGAGCCTCCTGATCTTA 1440  
QY 3812 TTGGGGGAGTATAAGACTGTTCCCAAAACCCATTGCACCTGGAGCCGCTTTTGGCAACA 3871  
Db 1441 TTGGGGGAGTATAAGACTGTTCCCAAAACCCATTGCACCTGGAGCCGCTTTTGGCAACA 1500  
QY 3872 AAGCCGCTGCTCTCTGTTGTTGTGAGGCTCCCTCGAGGCTGGGTGGGATCCCTCCTC 3931  
Db 1501 AAGCCGCTGCTCTCTGTTGTTGTGAGGCTCCCTCGAGGCTGGGTGGGATCCCTCCTC 1560  
QY 3932 CTGGGCAGTCAGGTCTTGCTGTGCGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 3991  
Db 1561 CTGGGCAGTCAGGTCTTGCTGTGCGCCAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1620  
QY 3992 TAGTGTAAGAGGAGAGTCAAGAGCCGTGAGAAACCGGAAGCAGAGCCCTGCAACAGC 4051  
Db 1621 TAGTGTAAGAGGAGAGTCAAGAGCCGTGAGAAACCGGAAGCAGAGCCCTGCAACAGC 1680  
QY 4052 CTGGGACCTGCATCTGA 4068  
Db 1681 CTGGGACCTGCATCTGA 1697

Search completed: March 5, 2004, 05:11:12  
Job time : 10247.8 secs



**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 4, 2004, 16:59:34 ; Search time 6445.51 Seconds  
(without alignments)  
18847.119 Million cell updates/sec

Title: US-09-787-097-11

Perfect score: 4068

Sequence: 1 atgttgccgcagcggcgcc.....agcctggacctgcattctga 4068

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```
EST: *
1: em_estba: *
2: em_esthum: *
3: em_estin: *
4: em_estnu: *
5: em_estov: *
6: em_estpl: *
7: em_estro: *
8: em_hc: *
9: gb_est1: *
10: gb_est2: *
11: gb_hc: *
12: gb_est3: *
13: gb_est4: *
14: gb_est5: *
15: em_estfun: *
16: em_estom: *
17: em_gss_hum: *
18: em_gss_inv: *
19: em_gss_pln: *
20: em_gss_vrt: *
21: em_gss_fun: *
22: em_gss_mam: *
23: em_gss_mus: *
24: em_gss_pro: *
25: em_gss_red: *
26: em_gss_phg: *
27: em_gss_vrl: *
28: gb_gss1: *
29: gb_gss2: *
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	3430.8	84.3	3976	29	AY418587 Homo sapi
2	2923.6	71.9	3976	29	AY418589 Mus muscu
3	2866	70.5	3671	29	AY418588 Pan trogl
4	925.6	22.8	1201	13	EX440935 EX440935

5	775	19.1	964	13	BUS06373
6	758.8	18.7	933	12	BG678679
7	699.6	17.2	840	14	CB520374
8	681	16.7	681	12	BM783739
9	680.4	16.7	683	12	BM783788
10	680	16.7	680	12	BM783258
11	680	16.7	680	12	BM783773
12	677	16.6	677	12	BM783718
13	674.4	16.6	677	12	BM783821
14	673	16.5	673	12	BM783685
15	665.4	16.4	792	14	CD644584
16	652.2	16.0	962	13	EX431593
17	636.8	15.7	674	12	BG563729
18	636.4	15.6	770	13	BY761276
19	636	15.6	777	14	CF745736
20	634	15.6	706	12	BI089332
21	632	15.5	632	12	BM773003
22	632	15.5	632	12	BM783264
23	628.2	15.4	953	10	BF313142
24	623	15.3	917	10	BF672370
25	618	15.2	633	10	BE894518
26	612.2	15.0	722	13	EX669806
27	608	14.9	735	13	BY761390
28	607.8	14.9	991	10	BE880905
29	599.2	14.7	940	13	BQ956223
30	589.6	14.5	743	14	CB519923
31	585.2	14.4	716	14	CF745901
32	580.2	14.3	714	10	BE650220
33	580	14.3	591	12	BI850289
34	569.8	14.0	653	14	CB446934
35	554.2	13.6	663	14	CF745835
36	549.6	13.5	902	13	BU151781
37	539	13.2	644	14	CB425680
38	537.4	13.2	942	10	BF316716
39	532.2	13.1	641	10	BB664616
40	524.2	12.9	626	10	BE980188
41	522.6	12.8	600	14	CB432517
42	517.6	12.7	595	10	BF853681
43	516.6	12.7	598	10	BF853006
44	513.2	12.6	544	10	BF853982
45	502.6	12.4	524	10	BE709733

#### ALIGNMENTS

RESULT 1  
AY418587  
LOCUS  
DEFINITION Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,  
Genomic survey sequence.  
ACCESSION AY418587  
VERSION AY418587.1 GI:39774547  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS  
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A.,  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 3976)  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
Inferring nonneutral evolution from human-chimp-mouse orthologous  
gene trios  
JOURNAL Science 302 (5652), 1960-1963 (2003)  
PUBMED 14671302  
REFERENCE  
AUTHORS  
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarimal, A.,  
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,  
Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,  
Adams, M.D. and Cargill, M.  
Direct Submission  
TITLE

3976 bp DNA linear GSS 17-DEC-2003

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,  
Rockville, MD 20850, USA  
COMMENT These sequences were made by sequencing genomic exons and ordering  
them based on alignment.  
FEATURES  
    source Location/Qualifiers  
        1..3976  
            /organism="Homo sapiens"  
            /mol\_type="genomic DNA"  
            /db\_xref="taxon:9606"  
            <1..>3976  
            /gene="ATRN"  
            /locus\_tag="HCM607"  
ORIGIN  
Query Match 84.3%; Score 3430.8; DB 29; Length 3976;  
Best Local Similarity 86.3%; Pred. No. 0;  
Matches 3432; Conservative 0; Mismatches 544; Indels 0; Gaps 0;  
Qy 93 TGTCAACGGCGTGTGCAACCTGGCACGGGCGAGTGGTCTGCCCGCGCGGTGGGT 152  
Db 1 TGTCAACGGCGTGTGCAACCTGGCACGGGCGAGTGGTCTGCCCGCGCGGTGGGT 60  
Qy 153 GGGGAGCAATGCCAGCACTGGCGGGCGCTTTCAGACTAATGGATCTTCGGGTTGT 212  
Db 61 GGGGAGCAATGCCAGCACTGGCGGGCGCTTTCAGNNNNNNNNNNNNNNNNNNNN 120  
Qy 213 GACAGATGGACCTGGAAATTAATAACAAACGAAAGTGCACTGGTCAATGAAGCA 272  
Db 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180  
Qy 273 GCCAATAGAAATATGAGACTCTGTTTCAATCAATTTGCTACAGAGTGTAGTTGGACCA 332  
Db 181 GCCAATAGAAATATGAGACTCTGTTTCAATCAATTTGCTACAGAGTGTAGTTGGACCA 240  
Qy 333 TTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCATTTAGTGGCT 392  
Db 241 TTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCATTTAGNNNNN 300  
Qy 393 CATGTCTCTGAGAGATGGCAATGAGACTGTCCCTGAGGTGTTGCCACATCAGTTA 452  
Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360  
Qy 453 TGCCTCTGCTCAATTTTATGATGCTGCTTAATAATTTGACTGGATTTAATATTACTTA 512  
Db 361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420  
Qy 513 CAGTTTGTATGTGTCCTCAATTAATCTGCTCAGCCGAGGAGGTGAAGTCAAGTAATAG 572  
Db 421 NNNNTTTGATGTGTCCTCAATTAATCTGCTCAGCCGAGGAGGTGAAGTCAAGTAATAG 480  
Qy 573 CAGGATACCTGTTGAATGTCAATGTTCTGAAACTGGAAGGTGAAGTGAACATTC 632  
Db 481 CAGGATACCTGTTGAATGTCAATGTTCTGAAACTGGAAGGTGAAGTGAACATTC 540  
Qy 633 TCACTGACAGCAACTGTGGTTTCTCTCATGAGGCATCTGCAATTCAGTGTGTCAG 692  
Db 541 TCACTGACAGCAACTGTGGTTTCTCTCATGAGGCATCTGCAATTCAGTGTGTCAG 600  
Qy 693 AGGATGCTCTGCTTCTCAGCTGGGAGGCTGCTGATGTTCAAGTCTCTGACAGCTAA 752  
Db 601 AGGATGCTCTGCTTCTCAGCTGGGAGGCTGCTGATGTTCAAGTCTCTGACAGCTAA 660  
Qy 753 CCAGTCAATTTGGACTCGAGAGGAATTTCTAACTTTAAAGCTCCCGAGAGCATCTCATAA 812  
Db 661 CCAGTCAATTTGGACTCGAGAGGAATTTCTAACTTTAAAGCTCCCGAGAGCATCTCATAA 720  
Qy 813 AGCTGTGCTCAATGGAAACATTTATGTTGGTGTGTTGGAGGATATATTTCAACCTCAGA 872  
Db 721 AGCTGTGCTCAATGGAAACATTTATGTTGGTGTGTTGGAGGATATATTTCAACCTCAGA 780  
Qy 873 TTATAACATGGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTTTC 932  
Db 781 TTATAACATGGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTTTC 840

Qy 933 TGTGAACAATGTGGTTGTTAGATATGGTCATCTTTGGCATTATACAGGATAAAATTTA 992  
Db 841 TGTGAACAATGTGGTTGTTAGATATGGTCATCTTTGGCATTATACAGGNNNNNNNNN 900  
Qy 993 CATGTATGGAGAAATTTGATTCAACTGGGAATGTGCAATGAGTTGAGAGTTTTC 1052  
Db 901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960  
Qy 1053 CATTCATAATCAGTCATGGGTGTTTGTGACCCCTTAAGCAAGAGCAGTATGCACTGT 1112  
Db 961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020  
Qy 1113 TGGGCATCTGCACACATTTGTTACACTGAAGAATGGCGAGTGGTCACTGTGTCATCTT 1172  
Db 1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080  
Qy 1173 TGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAAGATATGATTGGATAAGAA 1232  
Db 1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140  
Qy 1233 CACATGGAGTATATTACACACCCAGGGTGGCTTTGTGCAAGGGGGTTACGGCCATAGCAG 1292  
Db 1141 CACATGGAGTATATTACACACCCAGGGTGGCTTTGTGCAAGGGGGTTACGGCCATAGCAG 1200  
Qy 1293 TGTTTACACCATAGGACAGGGCCCTATAGCTTCATGTTGGTGGCTACAGGCTTTCAGTGC 1352  
Db 1201 TGTTTACACCATAGGACAGGGCCCTATAGCTTCATGTTGGTGGCTACAGGCTTTCAGTGC 1260  
Qy 1353 CAATAAGTACCGGCTTGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGGAC 1412  
Db 1261 CAATAAGTACCGGCTTGCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGGAC 1320  
Qy 1413 CATTTTAAAGACAGCGGATTTTCCGTACTTGGCAACAGCTGTGATGATGTGGAAC 1472  
Db 1321 CATTTTAAAGACAGCGGATTTTCCGTACTTGGCAACAGCTGTGATGATGTGGAAC 1380  
Qy 1473 CATCTGTGTTTGGGGGAAACACACAATGACACATCTATGAGCCATGGCGCAATG 1532  
Db 1381 CATCTGTGTTTGGGGGAAACACACAATGACACATCTATGAGCCATGGCGCAATG 1440  
Qy 1533 CTTCTCTCAGATTTCAATGCGCTATGATTCCTGTGACCGTGGTGTGTCAGTCTCCAG 1592  
Db 1441 CTTCTCTCAGATTTCAATGCGCTATGATTCCTGNNNNNNNNNNNNNNNNNNNNNN 1500  
Qy 1593 ACCTGATCTCCACCATGATGTCAACAGATTGGCCATTTCAGCAGTCTTACACACAGCAC 1652  
Db 1501 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1560  
Qy 1653 CATGTATGTGTTCCGGTGGTTTCAATAGTCTCTCTCCTCAGCGACATCTCTGTTATTCACCTC 1712  
Db 1561 CATGTATGTGTTCCGGTGGTTTCAATAGTCTCTCTCCTCAGCGACATCTCTGTTATTCACCTC 1620  
Qy 1713 GGAACAGTGTGATCGGATCGGAGTGAAGCGCTTGTAGCAGCAGGACTGGTATTCG 1772  
Db 1621 GGAACAGTGTGATCGGATCGGAGTGAAGCGCTTGTAGCAGCAGGACTGGTATTCG 1680  
Qy 1773 GTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACA 1832  
Db 1681 GTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACA 1740  
Qy 1833 AGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTGA 1892  
Db 1741 AGAAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTGA 1800  
Qy 1893 CCAGCACACAGATTTTACAGTGCACAGCCACACCAATGACTGCGCACTGGTGCATGA 1952  
Db 1801 CCAGCACACAGATTTTACAGTGCACAGCCACACCAATGACTGCGCACTGGTGCATGA 1860  
Qy 1953 CCATTGTGTCGCCAGGAACCAAGCTGCTCAGAGGCCAGATCTCCATTTTATAGGTATGA 2012  
Db 1861 CCATTGTGTCGCCAGGAACCAAGCTGCTCAGAGGCCAGATCTCCATTTTATAGGTATGA 1920

Qy	2013	GAATTGCCCCAAAGATAAACCTTATGTACTACTGTAAACAAGAACACAGCTGTCAGGAGCTG	2072
Db	1921	GAATTGCCCCAAAGATAAACCCATGTACTACTGTAAACAAGAACACAGCTGTCAGGAGCTG	1980
Qy	2073	TGCCCTGGAACCAAGACTGCCAGTGGGAGCCCCGGAATCAGAGTGTGAATGCCCTGCCCGA	2132
Db	1981	TGCCCTGGAACCAAGACTGCCAGTGGGAGCCCCGGAATCAGAGTGTGAATGCCCTGCCCGA	2040
Qy	2133	AAATATCTGTGGCAATTGGCTGGCAATTTGGTTGGAAACTCAATGTTGAAAATTAATACTGC	2192
Db	2041	AAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAACTCAATGTTGAAAATTAATACTGC	2100
Qy	2193	CAAGGAAATTATGACAAATGCTAAATTTGTTCTGTAGGAACCAAAATGCCCTTTTGCTTC	2252
Db	2101	CAAGGAAATTATGACAAATGCTAAATTTGTTCTGTAGGAACCAAAATGCCCTTTTGCTTC	2160
Qy	2253	TCATTACAACCCAGAAAGGTAGAAATTTGCTTCTTAAGCAGCTCGAATTAATCAAGTCAATC	2312
Db	2161	TCATTACAACCCAGAAAGGTAGAAATTTGCTTCTTAAGCAGCTCGAATTAATCAAGTCAATC	2220
Qy	2313	TCAGAGCATGTCNAAGTCAACCTTAACCCCATGGCTGGCTTCGGNAGATCAATGTGTC	2372
Db	2221	TCAGAGCATGTCNAAGTCAACCTTAACCCCATGGCTGGCTTCGGNAGATCAATGTGTC	2280
Qy	2373	CTACTGCTGTGGGAAGATATGTCCCCATTTACAATAGTTTACTACAGTGGATGCCGTC	2432
Db	2281	CTACTGCTGTGGGAAGATATGTCCCCATTTACAATAGTTTACTACAGTGGATGCCGTC	2340
Qy	2433	TGAGCCAGTGATGCTCGAATCTGTGGAAATTTATCAGAACCAAGTACTCGGGACTGAA	2492
Db	2341	TGAGCCAGTGATGCTCGAATCTGTGGAAATTTATCAGAACCAAGTACTCGGGACTGAA	2400
Qy	2493	GGCTGCAACCTGCAATCAACCACTCAATGTTAGTGTCTGTGAAAGGCTGCAAAACACAG	2552
Db	2401	GGCTGCAACCTGCAATCAACCACTCAATGTTAGTGTCTGTGAAAGGCTGCAAAACACAG	2460
Qy	2553	TGCTAAGCAGTGC CGGACACAATGTGCTTTGAGGACAGCATGTGGAGATTGCAACACGCG	2612
Db	2461	TGCTTAAAGCAGTGC CGGACACAATGTGCTTTGAGGACAGCATGTGGAGATTGCAACACGCG	2520
Qy	2613	CAGCTCTGAGTGCAATGTGGTCAGACACATGAGACAGTGTGGACTTCCAAATGCCATGT	2672
Db	2521	CAGCTCTGAGTGCAATGTGGTCAGACACATGAGACAGTGTGGACTTCCAAATGCCATGT	2580
Qy	2673	GGCTCTCTTCCCTTTGGCCAGTGATGGAATGGTATACGATGAGCACTGCCCCCTGA	2732
Db	2581	GGCTCTCTTCCCTTTGGCCAGTGATGGAATGGTATACGATGAGCACTGCCCCCTGA	2640
Qy	2733	AAATTGTTCAAGCTTACTGTACTGTAGTCATTGCTTGGAGCAACCAAGCTGTGGCTGGTG	2792
Db	2641	AAATTGTTCAAGCTTACTGTACTGTAGTCATTGCTTGGAGCAACCAAGCTGTGGCTGGTG	2700
Qy	2793	TACTGATCCCAAGCAATACTTGGCAAGGGAATGATAGAGGGTTTCTTATAAAGGACCACT	2852
Db	2701	TACTGATCCCAAGCAATACTTGGCAAGGGAATGATAGAGGGTTTCTTATAAAGGACCACT	2760
Qy	2853	GAAAGTGCCTTTCGAAGCCCTCAAGGAAATTTCTATCCACAGCCCTGTCAATTCAG	2912
Db	2761	GAAAGTGCCTTTCGAAGCCCTCAAGGAAATTTCTATCCACAGCCCTGTCAATTCAG	2820
Qy	2913	CATGTGCTAGAGGACAGCAGATACAACTGTGCTTTTCAATTCACTGTCCAGTGTGCCAATG	2972
Db	2821	CATGTGCTAGAGGACAGCAGATACAACTGTGCTTTTCAATTCACTGTCCAGTGTGCCAATG	2880
Qy	2973	CAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAACTTGACAC	3032
Db	2881	CAACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAACTTGACAC	2940
Qy	3033	AGGCAAGCACTGCAGGACCTGCATATCTGGCTTCTACGGTGATCCCAACCAATGGAGGAA	3092
Db	2941	AGGCAAGCACTGCAGGACCTGCATATCTGGCTTCTACGGTGATCCCAACCAATGGAGGAA	3000
Qy	3093	ATGTCAGCCATGCAAGTGCAATGGCGACCGCTCTCTGTGCAACCAACCAACGCGGCAAGTG	3152

Db	3001	ATGTGACCCATGCAAGTGCAAATGGGCACGCGTCTCTGTGCAACACCAACGAGGCAAGTG	3060
Qy	3153	CTTCTGCACCAACCAAGGCGCTCAAGGGGGACGAGTGCAGCTATGTGAGGTAGAAAAATCG	3212
Db	3061	CTTCTGCACCAACCAAGGCGCTCAAGGGGGACGAGTGCAGCTATGTGAGGTAGAAAAATCG	3120
Qy	3213	ATACCAAGGAAACCCCTCTCAGAGAAACATGTTATTATACCTTCTTCTTATGACATACAGTT	3272
Db	3121	ATACCAAGGAAACCCCTCTCAGAGAAACATGTTATTATACCTTCTTCTTATGACATACAGTT	3180
Qy	3273	CACCTTTAGTCTATCCACGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGCTAC	3332
Db	3181	CACCTTTAGTCTATCCACGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGCTAC	3240
Qy	3333	TCCTGACGAACAAACAGGATTTGGACATGTTTCATCAATGCCTCCAAAGAAATTTCAACCT	3392
Db	3241	TCCTGACGAACAAACAGGATTTGGACATGTTTCATCAATGCCTCCAAAGAAATTTCAACCT	3300
Qy	3393	CAACATCACCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTCGAGAAAGATGCCTGT	3452
Db	3301	CAACATCACCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTCGAGAAAGATGCCTGT	3360
Qy	3453	TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATCAGAAAGTTTGATTT	3512
Db	3361	TGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATCAGAAAGTTTGATTT	3420
Qy	3513	TCGCAACCAACCAATACACTTTCCTTGTGTTTATGTCAGTAATTTCACTCGGCCCATCAA	3572
Db	3421	TCGCAACCAACCAATACACTTTCCTTGTGTTTATGTCAGTAATTTCACTCGGCCCATCAA	3480
Qy	3573	AAATTCAGATGGCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGAC	3632
Db	3481	AAATTCAGATGGCTTCTCTCAGCACAGCAATTTTATGGACCTGGTACAGTTCTTCGTGAC	3540
Qy	3633	TTTCTTCAGTTGTTTCCCTCTTTGCTCTCGTGCTGCTGTTGGTGGAAAGATCAAAACA	3692
Db	3541	TTTCTTCAGTTGTTTCCCTCTTTTGTCTCTGGTGGCTGCTGTTGGTGGAAAGATCAAAACA	3600
Qy	3693	AAGTTGTTGGSCCTCCAGACGTACAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAG	3752
Db	3601	AAGTTGTTGGSCCTCCAGACGTACAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAG	3660
Qy	3753	CCGTCCCTTTGCCCTCTGTAAATGCGCTTGGAAACAGATGAGGAGCCTCTCTGATCTTAT	3812
Db	3661	CCGTCCCTTTGCCCTCTGTAAATGCGCTTGGAAACAGATGAGGAGCCTCTCTGATCTTAT	3720
Qy	3813	TGGGGGGAGTAAAGACTGTTCCCAACCCATTCGACTGAGGCGCTGTTTTTGGCAACAA	3872
Db	3721	TGGGGGGAGTAAAGACTGTTCCCAACCCATTCGACTGAGGCGCTGTTTTTGGCAACAA	3780
Qy	3873	AGCCGCTGCTCCTCTGTGTTTGTGAGGCTCCCTCGAGGCCGTGGGTGGCATCCCTCCCTCC	3932
Db	3781	AGCCGCTGCTCCTCTGTGTTTGTGAGGCTCCCTCGAGGCCGTGGGTGGCATCCCTCCCTCC	3840
Qy	3933	TGGGGAGTCAGGTCTGTTGGCCAGGCCCTGGTGGACATTTCTCAGCAGATGCCGAT	3992
Db	3841	TGGGGAGTCAGGTCTGTTGGCCAGGCCCTGGTGGACATTTCTCAGCAGATGCCGAT	3900
Qy	3993	AGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAAACCGGAAGCAGACGACCCCTGCACAGCC	4052
Db	3901	AGTGTACAAGGAGAAGTCAGGAGCCGTGAGAAAACCGGAAGCAGACGACCCCTGCACAGCC	3960
Qy	4053	TGGGACCTGCATCTGA 4068	
Db	3961	TGGGACCTGCATCTGA 3976	

## RESULT 2

AY418589

LOCUS

## DEFINITION

AY418589 3976 bp DNA linear GSS 17-DEC-2003  
Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,  
genomic survey sequence.

ACCESSION AY418589  
 VERSION AY418589.1 GI:39774549  
 KEYWORDS GSS.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 3976)  
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.  
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios  
 JOURNAL Science 302 (5652), 1960-1963 (2003)  
 PUBMED 14671302  
 REFERENCE 2 (bases 1 to 3976)  
 AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, P., Murphy, B., Ferreria, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J., Adams, M.D. and Cargill, M.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA  
 COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.  
 FEATURES  
 source 1..3976  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 <1..>3976  
 /gene="ATRN"  
 /locus\_tag="HCM6607"  
 ORIGIN  
 Query Match 71.9%; Score 2923.6; DB 29; Length 3976;  
 Best Local Similarity 78.3%; Pred. No. 0;  
 Matches 3115; Conservative 0; Mismatches 861; Indels 0; Gaps 0;  
 QY 93 TGTCAAGCGCGTGGCTGCAACCTGGCAGCGCGTGGTCTGCCCGCGCGTGGT 152  
 Db 1 TGTCAAGCGCGCGCTGGCAACCTGGCAGCGCGTGGTCTGCCCGCGCGTGGT 60  
 QY 153 GGGCGAGCAATGCGCAGCACTGCGGGGCGCGCTTCAGACTAATCGATCTTCTGGGTTGT 212  
 Db 61 GGGCGAGCAATGCGCAGCACTGCGGGGCGCGCTTCAGACTAATCGATCTTCTGGGTTGT 120  
 QY 213 GACAGATGGACCTGGAATATATAATACAAACGAAGTGCACGTGGCTCATTTGAAGGACA 272  
 Db 121 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 180  
 QY 273 GCCAAATAGATAATAGACTTCCTTTCAATCAATTTTGTCTACAGAGTGTAGTTGGGACCA 332  
 Db 181 GCCAAATAGATAATAGACTTCCTTTCAACCAATTTTGTCTACAGATGTAGCTGGGACCA 240  
 QY 333 TTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCATTTAGTGGCCT 392  
 Db 241 TTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCATTTAGTGGCCT 300  
 QY 393 CATTTGTTCTGAGAGATGCAATGAGCTGTCCCTGAGGTTGTTGCCACATCAGGTTA 452  
 Db 301 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 360  
 QY 453 TGCCTTGCTGCAATTTTATGATGCTGCTTATTAATTTGATCTGGAATTAATTAATCTTA 512  
 Db 361 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 420  
 QY 513 CAGTTTTGATATGTCCTCAATACCTGCTCAGCGCGAGAGTGTAAAGTCAAGTAAATAG 572  
 Db 421 NNNTTTGACATGTCCTCGAATTAATTTGCTCAGCGCGAGAGTGTAAAGTCAAGTAAATAG 480  
 QY 573 CAGCGATACACTGTTGAATGTGAATGTTCTGAAAACTCGAAAGGTGAAGCATGTGACATTC 632

Db 481 CAGCAGCGCTGTTGAGTGTGAATGTTCTGAAAACCTGGAAGGGGAGTCTGTTGACATTC 540  
 QY 633 TCACGTGACAGACAACTGTGGTTTTCTCATCGAGGCATCTGCAATCAAGTGTGTCAG 692  
 Db 541 TCACGTGACAGACAACTGTGGTTTTCTCATCGAGGCATCTGTAATGCAAGCATACCA 600  
 QY 693 AGGATGCTCTGCTTCTCAGACTGCGAGGCTCTGGATGTTGAGTCTCTGTTACAGGATA 752  
 Db 601 AGGGTGCTCTGCTTCTCAGACTGCGAGGCTCTGGATGTTGAGTCTCTGTTGCGAGCTAA 660  
 QY 753 CCAGTCATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCGCAGAGCATCTCTATA 812  
 Db 661 CCAGTCATTTTGGACTCGAGAGGAATATTTCTGAATTAAGCTCCCGAGAGCTCTCTATA 720  
 QY 813 AGCTGTGTCATGGAACATTTATGTGGTTGTTGGAGGATATATGTTCAACCATCAGA 872  
 Db 721 AGCTGTGTCATGGAACATTTATGTTGGTTGTTGGCGGATATATGTTCAACCATCAGA 780  
 QY 873 TTATAACATGGTTCTAGGCTATGACCTTGTCTTAGGGAGTGGCTCCACATAAACCGTTC 932  
 Db 781 TTACAGCATGTTTGTAGGCTATGACCTTCTTAGGAAATGGCTTCACTTAACCATTC 840  
 QY 933 TGTGAACAATGTTGTTGATATGTCATCTTTGGCATATATACAGAGATAAAATTTA 992  
 Db 841 TGTGAACAATGTTGTTGATATGTCATCTTTGGCATATATACAGAGATAAAATTTA 900  
 QY 993 CATGTATGAGGAAAAATTTGATTCACCTGGGAATGTGACCAATCAGTTGAGAGTTTCA 1052  
 Db 901 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 960  
 QY 1053 CATTCATAATGAGTCATGGGTGTTGTCACCCCTTAAGCAAGAGCAGTAGTCAGTGGT 1112  
 Db 961 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1020  
 QY 1113 TGGGCACCTGTCACACATTTGTTACCTGAAGAAATGGCGAGTGGTCTGCTGGTCACTTT 1172  
 Db 1021 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1080  
 QY 1173 TGGTCACCTGCTCTCTATGATATATAAGCAATATGTCAGGAATATGATTTGGATAAGAA 1232  
 Db 1081 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1140  
 QY 1233 CACATGAGATATATTAACACCCAGGGTCCCTTGTGCAAGGGGGTTACGGCCATAGCAG 1292  
 Db 1141 CACATGAGATATATTAACATCTACAGGGTCTCTTGTGCAAGGGGGTTATGGCCACAGTAG 1200  
 QY 1293 TGTTTACGACCATAGGACCGGCGCTTATAGCTTTCATGTTGGCTTACAGGGCTTTTCAGTGC 1352  
 Db 1201 TGTTTATGATGACAGGACCAAGGCTCTGTAGCTTCAAGGCTTCAAGGCTTTCAGCGC 1260  
 QY 1353 CAATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGATACCCAGATGTGGAC 1412  
 Db 1261 CAACAAAATACCGGCTTGAGATGATGATCTCTACAGATGATGATGTGATGTGATGTGAC 1320  
 QY 1413 CATTCCTTAAGCAGACCGGATTTTCCGTTACTTGCACACAGCTGTGATGTAGTGTGGAAC 1472  
 Db 1321 CATTCCTTAAGCAGACCGGATTTTCCGTTACTTGCATACAGCTGTGATGTAGTGTGGAAC 1380  
 QY 1473 CATGCTGTTGTTGGGGGAAACACACAATGACATCTATAGCCATGCGCGCAATG 1532  
 Db 1381 CATGCTGTTGTTGGAGGAAACACACAATGACATCTTCCATGAGCCACGCTGCGCAATG 1440  
 QY 1533 CTTCTCTTCAGATTTTCATGGCTATGATGCTGTTGACCGCTGGTCTGCTGCTTCCAG 1592  
 Db 1441 CTTCTCTTCAGATTTTCATGGCTATGATGCTTATGATGCTTATGATGCTTATGATGCTT 1500  
 QY 1593 ACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGCTTTACACACAGCAC 1652  
 Db 1501 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1560  
 QY 1653 CATGTATGTTGCTGGTGGTTCAATAGTCTCTCTCAGCGACATCTGTTTACCTTC 1712

Db 1561 CATGTATGTCTCGGGGCTTCAACAGCCCTCTCTCCTCAGTGAGCTCTTGGTCTTTACCTC 1620  
Qy 1713 GGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTAGCAGAGGACCTGATTCG 1772  
Db 1621 GAGCAGTGTGATGCAACCGCATGAAGCTGCTTGTGTGGCAGAGGACCTGGTATCCG 1680  
Qy 1773 GTGTGTGTGAACACAGGCTGCTCTCAGTGTATCTCTGTGGGCTGTGGCACTGATGAACA 1832  
Db 1681 GTGTGTGTGGACACACAGTGTCTCTGATGTACCTCTCTGGAGTTGGCAACTGAAGACA 1740  
Qy 1833 AGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGA 1892  
Db 1741 AGCAGAAAGTTAAATCAGAGTGTGTTTTCTAAAGAAACCTTTGACCATGACAGATGTGA 1800  
Qy 1893 CCAGCACACAGATGTTTACAGCTGACAGCCACCAATGACTGTGCACTGTGTGAATGA 1952  
Db 1801 CCAGCACACAGATGTTTACAGCTGACAGCCCAATGCAATGACTGTGCACTGTGTGAATGA 1860  
Qy 1953 CCATTTGTCTCCAGGAAACACAGCTGCTCAGAGGCCAGATCTCCATTTTTAGGTATGA 2012  
Db 1861 TCAGTGTCTCCCTGTGAACACAGCTGCAAGAGGCCAGATCTCCATTTGCCAATATGA 1920  
Qy 2013 GAATTCGCCCAAGGATAACCTATGTACTACTGTAAAGAACCAAGAGCCAGCTGCAGGAGCTG 2072  
Db 1921 GAGTTGCCCAAGGATAACCCCATGTACTACTGCAATAGAAACCAAGCTGCAGGAGCTG 1980  
Qy 2073 TGCCTGTGACCAAGACTGCCAGTGGAGGCCCGGGAATCAGGAGTGCAATGCCCTGCCGA 2132  
Db 1981 TGCCCTAGACCAAGACTGCCAGTGGAGGCCCGGGAATCAAGAGTGCAATGCCCTGCCGA 2040  
Qy 2133 AAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAACCAATCATGTTTGAATAATTACTACTG 2192  
Db 2041 AAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAACCAATCATGTTTGAATAATTACTACTG 2100  
Qy 2193 CAAGGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAATGCTTGGCTTC 2252  
Db 2101 TAAGGAGAAATATGACAAATGCTAAATGTTCTGTAGGAACCAATGCTTGGCTTC 2160  
Qy 2253 TCTTACAAACCAAGAGGATGAATTTGCTTTAAGAGCTGCGAATTAATGAGTATC 2312  
Db 2161 CCTCACATCCAGAGAGGAGTGTGCTTAAAGCAGCTTCGATTAATGCAATCATC 2220  
Qy 2313 TCAGAGCATCTCAAGCTCACTTAACCCCATGGTGGCTTCGGAAGATCAATGTGC 2372  
Db 2221 TCAAGATATGTCAGAGCTCACTCTGACTCATGGGTTGGTCTTCGGAAGATCAATGTGC 2280  
Qy 2373 CTACTGTGTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCCGC 2432  
Db 2281 TTACTGTGTGGGAGGATATGCTCAATTCACAAATAGTTTGTCTGCAAGTGCATGCCATC 2340  
Qy 2433 TGAGCCAGTGTGCTGGATCTGTGGAATTTTATCAGAACCCAGTACTCGGGAGCTGAA 2492  
Db 2341 TGAGCCAGTGTGCTGGCTTCTGTGGGATCTTGTGAGAGCTTAGTACTCGGGGATTA 2400  
Qy 2493 GGCTGCAACCTGCATCAACCCATCAATGGTAGTGTGTGAAGCCCTGCAACCCACAG 2552  
Db 2401 GGCTGCAACCTGCATCAACCCCTCAATGCGAGCTGTGTGAAGCCCTGCAACCCACAG 2460  
Qy 2553 TGTAGCAGTGTGGCAGCACCATGTGCTTGAGGACAGCATGTGGAGATGCAACAGCGG 2612  
Db 2461 TGCCCAAGCAGTGTGGGACACCATGTGCCCTGCGGACAGCGTGTGGCAGTGCATAGCAG 2520  
Qy 2613 CAGCTCTGAGTGTGCTGGTGCAGCAACATGAAGCAGTGTGTGGAATCCCAATGCCCTATGT 2672  
Db 2521 CAGCTCGAGTGTGATGTGTGAGTAAATGATGAAGCAGTGTGTGGAATCCCAATGCCCTATGT 2580  
Qy 2673 GGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTTGCCCCCTGA 2732  
Db 2581 GGCTCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTTGCCCCCTGA 2640  
Qy 2733 AAATTTGTCAGGCTACTGTACTGTAGTCAATTCCTTGGAGCAACAGGCTGTGGCTGTG 2792  
Db 2641 AAATTTGTCAGGCTACTGTACTGTAGTCAATTCCTTGGAGCAGCCAGGCTGTGGTGGTG 2700

Qy 2793 TACTGATCCAGCAATACTGTGCAAGGAAATGATAGAGGGTTCTCTATAAGGACACAGT 2852  
Db 2701 TACTGATCCATAGCAATACTGTGCAAGGAAATGATAGAGGACAGCTATATAAGGACCTGT 2760  
Qy 2853 GAAAGTGCCTTGGCAAGCCCTTATCCAGGAAATTTCTATCCAGGCCCCCTGCTCAATTCAG 2912  
Db 2761 GAAGATGCCGTCAAGGCTCTGTGAGGAAATGTGTATCCACAGGCCCTTCTGAACCTCCAG 2820  
Qy 2913 CATGTGTCTAGAGGACAGCAGATACAACTGTCTTTTCTATTCACCTGTCCAGCTTCCCAATG 2972  
Db 2821 CATGTGTCTAGAGGACAGCAGATACAACTGTCTTTTCTATTCACCTGTCCAGCTTCCCAATG 2880  
Qy 2973 CAAAGCCCAAGTAAATGCAATCAGAGCATCTGTGAGAAATGTGAGAACTCTGAGCAAC 3032  
Db 2881 CAAAGCCCAAGTAAATGCAATCAGAGCATCTGTGAGAAATGTGAGAACTCTGAGCAAC 2940  
Qy 3033 AGGCAAGCAGTGCAGAGCTGCTATCTGGCTTTCTACGGTGTATCCACCAATGAGGGAA 3092  
Db 2941 GGGCAAGCAGTGCAGAGCTGCTATCTGGCTTTCTATGGTGTACCCGACTTAATGAGGCAA 3000  
Qy 3093 ATGTGAGCCATGCAAGTGGCAATGGGCGCTCTCTGTGCAACCAACACCGGCAAGTG 3152  
Db 3001 ATGTGAGCCATGCAAGTGGCAATGGGCGCTCTCTGTGCAACCAACACCGGCAAGTG 3060  
Qy 3153 CTCTGTGCAACCAAGGCGCTCAAGGGGCAAGAGTGCAGCTATGTGAGGTAGAAAATCG 3212  
Db 3061 CTCTGTGCAACCAAGGCGCTCAAGGGGCAAGAGTGCAGCTATGTGAGGTAGAAAATCG 3120  
Qy 3213 ATACCAAGAAACCTCTCAGAGGAAATGTTATTTATCTCTCTTATTTGACTATCAGTT 3272  
Db 3121 ATACCAAGAAACCTCTCAGAGGAAATGTTATTTATCTCTCTTATTTGACTATCAGTT 3180  
Qy 3273 CACCTTTAGTCTATCCAGGAAAGATGCTGCTATTTACAGCTATCAATTTTGGCTPAC 3332  
Db 3181 CACCTTTAGTCTATCCAGGAAAGAGCGCTCTACTACAGCCATCAACTTTTGGCTPAC 3240  
Qy 3333 TCCTGCAACCAAGGAGTGTGGAGATTTGAGATGTTTCAATGCTCTCAGAAATTTCAACCT 3392  
Db 3241 TCCTGATGAACAAACAGGATTTGAGATGTTTCAATGCTCTCAGAAATTTCAACCT 3300  
Qy 3393 CAAACATCCTGCGGTGCGCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGTGCCTGT 3452  
Db 3301 CAAACATCCTGCGGTGCGCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGTGCCTGT 3360  
Qy 3453 TGTTCAAAAACCAACATTAAGAGATCAAAAGATGATTTCTCTAATGAGAGTTGATTT 3512  
Db 3361 TGTTCAAAAACCAACATTAAGAGATCAAAAGATGATTTCTCTAATGAGAGTTGATTT 3420  
Qy 3513 TCGCAACCAACCAACATCACTTTCTTTTGTATGTGAGTAATTTTCACTGCGCCCATCAA 3572  
Db 3421 TCGCAACCAACCAACATCACTTTCTTTTGTATGTGAGTAATTTTCACTGCGCCCATCAA 3480  
Qy 3573 AATTGAGATTGCTTCTCTCAGCAAGCAATTTTATGSAACCTGTGTACAGTTTCTTGTGAC 3632  
Db 3481 AATTGAGATTGCTTCTCTCAGCAAGCAATTTTATGSAACCTGTGTACAGTTTCTTGTGAC 3540  
Qy 3633 TTTCTTCAAGTTTCTCTCTCTTTGCTCTGTGCTGTGCTGTGTTGGAAGATCAACA 3692  
Db 3541 TTTCTTCAAGTTTCTCTCTCTCTCTCTGTGCTGTGCTGTGTTGGAAGATCAACA 3600  
Qy 3693 AAGTTGTTGGGCTCTCAGACGTAGAGGCAACTTTCTCGAGAGATGCAACAGATGGCCAG 3752  
Db 3601 GAGCTGTTGGGCTCTCAGGCGGAGAGGCAACTTTCTCGGAGATGCAACAGATGGCCAG 3660  
Qy 3753 CGGTCCTTGTGCTCTGTAAATGTGCGCTTGGAAACAGATGAGGAGCTCTCTGATCTAT 3812  
Db 3661 CGGTCCTTGTGCTCTGTAAACGTGTGCTTGGAAACAGATGAGGAGCTCTCTGATCTAT 3720  
Qy 3813 TGGGGGAGTATAAGAGCTGTTCCCAACCATTTGCACTGGAGCGCTGTTTGTGGCAACA 3872  
Db 3721 TGGGGGAGTATAAGAGCTGTTCCTAAGCCCATTTGCGCTTGGAGCCTCTTGTGGTAACA 3780







Db	3421	AAAGTTGTTGGGCTCCAGAGCTAGAGCAACTTCTCGAGAGATGCAACAGATGGCCAG	3480
Qy	3753	CCGTCCTCTTGGCTCTCTGTAATGTCGCTTTGGAACACAGATGAGGAGCTCTCTGATCTTAT	3812
Db	3481	CCGTCCTCTTGGCTCTCTGTAATGTCGCTTTGGAACACAGATGAGGAGCTCTCTGATCTTAT	3540
Qy	3813	TGGGGGAGTATAAAGACTGTTCCCAACCCATGCACTGGAGCCGTTGTTGGCAACAA	3872
Db	3541	TGGGGGAGTATAAAGACTGTTCCCAACCCATGCACTGGAGCCATGTTTGGCAACAA	3600
Qy	3873	AGCGCTGCTCTCTCTGTTGAGGCTCCCTCGAGGCTGGGATCCCTCTCTCC	3932
Db	3601	AGCGCTGCTCTCTCTGTTGAGGCTCCCTCGAGGCTGGGATCCCTCTCTCC	3660
Qy	3933	TGGGCACTCAG	3943
Db	3661	TGGGCACTCAG	3671
RESULT 4			
BX440935			
LOCUS			
DEFINITION BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone			
CS0DP012YL19 5-PRIME, mRNA sequence.			
ACCESSION BX440935			
VERSION BX440935.1 GI:30781889			
KEYWORDS EST.			
SOURCE Homo sapiens			
ORGANISM Homo sapiens (human)			
REFERENCE			
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
TITLE Li, W.B., Gruber, C., Jessee, J., and Polayes, D.			
JOURNAL Full-length cDNA libraries and normalization			
COMMENT Unpublished (2001)			
Contact: Genoscope			
Genoscope - Centre National de Sequencage			
BP 191 91006 EVRY cedex - France			
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr			
Library was constructed by life technologies, a division of			
Invitrogen. This sequence belongs to sequence cluster 10212.f For			
more information about this cluster, see			
http://www.genoscope.cns.fr/			
cgi-bin/cluster.cgi?seq=CS0DP012CF10QPI&cluster=10212.f. Contact :			
Feng Liang Email : fliang@lifetech.com URL :			
http://fulllength.invitrogen.com/ invitrogen Corporation 1600			
Paraday Avenue Genoscope sequence ID : CS0DP012CF10QPI.			
Location/Qualifiers			
1. 1201			
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/clone="CS0DP012YL19"			
/tissue_type="FETAL BRAIN"			
/dev_stage="fetal"			
/clone_lib="Homo sapiens FETAL BRAIN"			
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA			
was primed with a NotI-oligo(dT) primer. Five prime end			
enriched, double-strand cDNA was digested with Not I and			
cloned into the Not I and EcoRV sites of the pCMVSPORT 6			
vector. Library was not normalized."			
FEATURES			
source			
ORIGIN			
Query Match			
Best Local Similarity 22.8%; Score 925.6; DB 13; Length 1201;			
Matches 962; Conservative 4; Mismatches 8; Indels 4; Gaps 3;			
Qy	2115	GTGATTCGCTCCCGCAAAATATCTGTGGCATGGCTGGCATTTGGTTGGAACTCATG	2174
Db	57	GTGATTCGCTCCCGCAAAATATCTGTGGCATGGCTGGCATTTGGTTGGAACTCATG	116
Qy	2175	TTTCAAAATTACTCTCCAGAGAGATTTATGCAATGCTAAATTTGTTGTAGGAACCA	2234

Db	117	TTTGAATAATTACTACTGCCAAGGAGAAATTATGCAATGCTAAATTTGTTGTAGGAACCA	176
Qy	2235	CAATGCCCTTTTGGCTTCTCTTTACAACCCAGAAAGGTAGAAATTTGTCCTTAAGCAGCT	2294
Db	177	CAATGCCCTTTTGGCTTCTCTTTACAACCCAGAAAGGTAGAAATTTGTCCTTAAGCAGCT	236
Qy	2295	CGAATAATGCAATCACTCAGAGCATGTCCAGCTCACCTTAACCCCATGGTTCGGCT	2354
Db	237	CGAATAATGCAATCACTCAGAGCATGTCCAGCTCACCTTAACCCCATGGTTCGGCT	296
Qy	2355	TCGGAAGATCAATGTGCTACTGCTGCGAAGATATGTCCTCCCATTTTACAATAGTTT	2414
Db	297	TCGGAAGATCAATGTGCTACTGCTGCGAAGATATGTCCTCCCATTTTACAATAGTTT	356
Qy	2415	ACTACAGTGGATGCGCTGAGCCAGTGTGATGCTGCGAATTTTATCAGAAC	2474
Db	357	ACTACAGTGGATGCGCTGAGCCAGTGTGATGCTGCGAATTTTATCAGAAC	416
Qy	2475	CAGTACTCGGGACTGAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTCTCTGA	2534
Db	417	CAGTACTCGGGACTGAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTCTCTGA	476
Qy	2535	AAGGCTCTCAAAACACACAGTCTAAGCAGTGTGCGGACCACTGTCCTTGAGGACGATG	2594
Db	477	AAGGCTCTCAAAACACACAGTCTAAGCAGTGTGCGGACCACTGTCCTTGAGGACGATG	536
Qy	2595	TGGAGATTGACACGCGGAGCTCTGAGTGTGCTGCGGACCACTGAGCAGTGTG	2654
Db	537	TGGAGATTGACACGCGGAGCTCTGAGTGTGCTGCGGACCACTGAGCAGTGTG	596
Qy	2655	GGACTCCCAATGCTATGCGCTCTCTTCCCTTTTGGCCAGTGTGATGAATGGTATACGAT	2714
Db	597	GGACTCCCAATGCTATGCGCTCTCTTCCCTTTTGGCCAGTGTGATGAATGGTATACGAT	656
Qy	2715	GAGCAGCTGCCCCCTGAAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCA	2774
Db	657	GAGCAGCTGCCCCCTGAAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCA	716
Qy	2775	ACGAGGCTGTGCTGCTGTACTGATCCAGCAATCTGCGCAAGGAAATGCTATAGAGG	2834
Db	717	ACGAGGCTGTGCTGCTGTACTGATCCAGCAATCTGCGCAAGGAAATGCTATAGAGG	776
Qy	2835	TTCTATAAAGACACAGTGAAGTGCCTTCGCAAGCCCTACAGGAAATTTCTATCCACA	2894
Db	777	TTCTATAAAGACACAGTGAAGTGCCTTCGCAAGCCCTACAGGAAATTTCTATCCACA	836
Qy	2895	GCCCTGCTCAATTCAGCATGTCTAGAGGACAGCAGATACAACTGCTCTTCAATTC	2954
Db	837	GCCCTGCTCAATTCAGCATGTCTAGAGGACAGCAGATACAACTGCTCTTCAATTC	896
Qy	2955	CTGTCCAGCTTTCGCAATGCAACCGCCACAGTAAATGCAATCAGAGCATCTGTGAA	3014
Db	897	CTGTCCAGCTTTCGCAATGCAACCGCCACAGTAAATGCAATCAGAGCATCTGTGAA	956
Qy	3015	GTGTGAGAACTTGACACAGCAGCAGTGTGAGGACAGCAGATACAACTGCTTACGCTGA	3074
Db	957	GTGTGAGAACTTGACACAGCAGCAGTGTGAGGACAGCAGATACAACTGCTTACGCTGA	1013
Qy	3075	TCCACCAATGGAGGAA 3092	
Db	1014	T-CCACCAATGGAGGAA 1030	
RESULT 5			
BU506373			
LOCUS			
DEFINITION BU506373 964 bp mRNA linear EST 12-SEP-2002			
AGENCOURT 10015485 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6494441			
5', mRNA sequence.			
ACCESSION BU506373			
VERSION BU506373.1 GI:22812606			
KEYWORDS EST.			
SOURCE Mus musculus (house mouse)			
ORGANISM Mus musculus			

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 964)

REFERENCE  
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov

Tissue Procurement: The Cepko Laboratory  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

Plate: L1AM14050 row: k column: 18

High quality sequence stop: 592.

FEATURES

source

Location/Qualifiers

1..964

/organism="Mus musculus"

/mol\_type="mRNA"

/db\_xref="taxon:10090"

/clone="IMAGE:649441"

/tissue\_type="retina"

/lab\_host="DH10B (phage-resistant)"

/clone\_lib="NIH MGC 94"

/note="Organ: eye; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dr primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 19.1%; Score 775; DB 13; Length 964;

Best Local Similarity 88.9%; Pred. No. 7.8e-186;

Matches 857; Conservative 0; Mismatches 105; Indels 2; Gaps 2;

QY 1774 TGTGTGGACACAGGTCGTCTCAGTGTAUUCGTGGCGGTGGCAACTGATGAACAA 1833

DB 1 TGTCTGTGGACACACAGTCGTCTCGATGTACCTCTCTGGAGTGGCAACTGAAGAACAA 60

QY 1834 GAAGAAAGTTAAATCAGAACTGTTTCCAAAGAACTCTTGACCATCAGACATGTGAC 1893

DB 61 GCAGAAAGTTAAATCAGAGTGTTTCTAAAGAAACCTTGACCATCAGACATGTGAC 120

QY 1894 CAGCACACAGATTGTTACAGCTGCACAGCCACACCAATGACTGCCACTGGTGCATGAC 1953

DB 121 CAGCACACAGATTGTTACAGCTGCACAGCCAAATACCAATGACTGCCACTGGTGCATGAT 180

QY 1954 CATTGTCTCCAGGACACACAGCTGCTCAGAGGCGCAGATCTCCATTTTGTAGTATGAG 2013

DB 181 CATTGTCTCCCTGTGAACACAGCTGCACAGAGGCGCAGATCTCCATTTCCCAAGTATGAG 240

QY 2014 AATTGCCCAAGGATAACCCCTATGTACTCTGTAACAAGAACACCACTGCAGGAGCTGT 2073

DB 241 AGTTGCCCAAGGATAACCCCTATGTACTCTGTAACAAGAACACCACTGCAGGAGCTGT 300

QY 2074 GCCTTGACACAGAACTGCAGTGGGAGCCCGGGAATCAGAGTGCAATTCCTCTGCCGAA 2133

DB 301 GCCTTAGACAGAACTGCCAGTGGGAGCCCGGGAATCAAGAGTGCAATTCCTCTGCCGAA 360

QY 2134 AATATCTGTGGCAATTTGGCTGGCAATTTGGTGGAAATCTCATTTTGAATACTACTCTGC 2193

DB 361 AATATCTGTGGCAATTTGGCTGGCAATTTGGTGGAAATCTCTGTCTGAAATCACTACTGTCT 420

QY 2194 AAGGAGAAATATGCAATGCTAAATTTGTTCTGTAGGAACCAACAATCCCTTTTGGTTCT 2253

DB 421 AAGGAGAAATATGCAATGCTAAATTTGTTCTGTAGGAACCAACAATCCCTTTTGGTTCT 480

QY 2254 CTTACACCCAGAGAGGTAGATTGTCCTTTAAGCAGCTGCATATGAGTCATCT 2313

DB 481 CTCACATCCAGAGAGGTGGAGTTGTCCTTTAAGCAGCTGCATATGAGTCATCT 540

QY 2314 CAGAGCATGTCCAAAGCTCACCTTAACCCCTGAGGCTGGCCCTCGAAGATCAATGTGTCC 2373

DB 541 CAAGATATGTCCAAAGCTCACTCTGACTCCATGGGTGGTCTTCGGAAGATCAATGTGTCT 600

QY 2374 TACTGTGTCTGGGAAGATATGTCCCCATTTTACAAATAGTTTACTACAGTGGATCCGTCT 2433

DB 601 TACTGTGTCTGGGAAGATATGTCTCCATTACAAATAGTTTGTCTGAGTGGATGCCATCT 660

QY 2434 GAGCCCACTGATGTGGATTCTGTGGATTTTATCAGAACCCAGTACTC-CGGGACTGAA 2492

DB 661 GAGCCCACTGATGTGGATTCTGTGGATTTTATCAGAGCTGTACTTCGGGGGATTAATA 720

QY 2493 GGCTGCACACCTGCATCAACCCCACTCAATGGTAGTGTCTGTGAAGAGGCTGCAAAACCAAG 2552

DB 721 GGCTGCACACCTGCATCAACCCCTCTCAATGGCAGAGCTGTGTGAAGAGGCTGCAAAACCAAG 780

QY 2553 TGCTAAGCAGTGGCGGACACCATGTGCTTGGAGACAG-CATGTGGAGATTGACACAGCG 2611

DB 781 TGCCCAAGCAGTGGCGGACACCATGTGCTTGGAGACAGCTTGTGGCAGTGGCACTAGCA 840

QY 2612 GCAGCTCTGAGTGCATGTGTGGCAGCAACATGAAGCAGTGTGTGAGCTTCCCAATGCTATG 2671

DB 841 CGAGCTCGAGTGCATGTGTGGCAGTACATGAAGCAGTGTGGGACTTCCCAATGCTATG 900

QY 2672 TGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATPACGATGAGCACTGCCCCCTG 2731

DB 901 TGGCTCTCTCCCTTTTGGCCAGTGTATGGAAGTATACGATGAGCAGCTGTCCAACTG 960

QY 2732 AAAA 2735

DB 961 AAAA 964

RESULT 6

BG678679

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BG678679 933 bp mRNA linear EST 01-MAY-2001  
602624477F1 NCI\_QGAP\_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',  
mRNA sequence.

ACCESSION  
BG678679  
VERSION  
BG678679.1 GI:13910076  
KEYWORDS  
EST.

SOURCE  
Homo sapiens (human)

ORGANISM  
Homo sapiens (human)

REFERENCE  
1 (bases 1 to 933)  
NIH-MGC <http://mgc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)

AUTHORS  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs@mail.nih.gov

TITLE  
Tissue Procurement: James Cleaver, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

COMMENT  
DNA sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>

PLATE: L1AM10602 row: j column: 20  
High quality sequence stop: 762.

FEATURES  
Location/Qualifiers

1..933  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4749427"  
/tissue\_type="squamous cell carcinoma"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_lib="NCI\_QGAP\_Skn4"

/note="Organ: Skin; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally. Primer: Oligo dr. Average insert size 1.5kb. Library constructed by Life Technologies. Note: this is a NCI\_QGAP Library."

```

FEATURES
  source
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
    1. (bases 1 to 840)
    NIH-MGC http://mgc.nci.nih.gov/.
    National Institutes of Health, Mammalian Gene Collection (MGC)
    Unpublished (1999)
    Contact: Robert Strausberg, Ph.D.
    Email: cgabbs@mail.nih.gov
    Tissue procurement: Dr. Jim Lin, University of Iowa
    cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
    cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
    DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
    Clone Distribution: Distribution information can be found at
    http://genome.uiowa.edu/distribution/mousefl.html
    This clone was contributed by the Brain Molecular Anatomy Project
    (BMAP)
    Seq primer: pYX-5.
    Location/Qualifiers
      1..840
        /organism="Mus musculus"
        /mol_type="mRNA"
        /strain="C57BL/6"
        /db_xref="taxon:10090"
        /clone="IMAGE:6840067"
        /tissue_type="whole brain"
        /dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
        /lab_host="DH10B (T1 phage resistant)"
        /clone_lib="NIH BMAP G10"
        /notes="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
        Site: Not I; The library was constructed according
        Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
        1996. Denatured RNA was size fractionated on a 1% agarose
        gel. First strand cDNA synthesis was primed with oligo-dT
        primer containing a Not I site. Double strand cDNA was
        size selected according to mRNA size fraction, ligated
        with EcoR I adaptor, digested with NotI and then cloned
        directionally into pYX-Asc vector. The library tag
        sequence located between the Not I site and the polyA tail
        is AGGAGACAGC. This library was created for the University
        of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
        Developing Mouse Nervous System', supported by National
        Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
        program coordinator."

```

## ORIGIN

Query Match	17.2%	Score 699.6	DB 14	Length 840
Best Local Similarity	90.1%	Pred. No. 1.2e-166		
Matches 758	Conservative 0	Mismatches 82	Indels 1	Gaps 1
QY	1929	CAATGACTGCCACCTGGTGCATGACCAATGTTGCCCCAGGAACCAACAGCTGCTCAGAAGG	1988	
DB	1	CAATGACTGCCACCTGGTGCATGACCAATGTTGCCCCAGGAACCAACAGCTGCAACAAGG	60	
QY	1989	CCAGATCTCCATTTTATAGTATGAGAAATGCCCCCAAGGATAACCTATGTACTACTGTAA	2048	
DB	61	CCAGATCTCCATTTGCAAGTATGAGAGTTGCCCCAGGATAACCCCATGTACTACTGCAA	120	
QY	2049	CAAGAAGACCAAGCTGCAGGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGCCCCGGAA	2108	
DB	121	TAAGAAACCAAGCTGCAGGAGCTGTGCCCTAGACCAGAACTGCCAGTGGGAGCCCCGGAA	180	
QY	2109	TCAGGAGTGCAATGCCCTGGCCCCGAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGAAA	2168	
DB	181	TCAAGAGTGCAATGCCCTGGCCCCGAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGAAA	240	
QY	2169	CTCATGTTTGAATAATTACTACTGCCAGGAGAAATTATGCAAACTGCTAAATTTGCTGTAG	2228	
DB	241	CTCGTGCTGAAAATCACTACTGCTAAGGAGAAATTATGCAAACTGCTAAATTTGCTGTAG	300	
QY	2229	GAACCACAATGCCCTTTTGGCTTCTCTAACACCAGGAAGTAGAAATTTGTCCTTAA	2288	
DB	301	GAACCACAATGCCCTTTTGGCTTCCCTCACATCCAGGAAGGTGAGTTTGTCTTAA	360	

ORIGIN	Query Match	18.7%; Score 758.8; DB 12; Length 933;
	Best Local Similarity 98.4%; Pred. No. 1e-181;	
	Matches 777; Conservative 0; Mismatches 12; Indels 1; Gaps 1;	
QY	2775	ACCAGGCTGGCTGGTGTACTGATCCGACATATCTGCCAAGGGAATGCTATCAGGG 2834
DB	2	ACCAGGCTGGCTGGTGTACTGATCCGACATATCTGCCAAGGGAATGCTATCAGGG 61
QY	2835	TTCTTATAAGGACCACTGGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCACA 2894
DB	62	TTCTTATAAGGACCACTGGAAGATGCTTCGCAAGCCCTTACAGGAAATTTCTATCCACA 121
QY	2895	GCCCTGCTCAATTCAGACATGTGTAGAGGACAGAGATACAACTGGTCTTTCAATCA 2954
DB	122	GCCCTGCTCAATTCAGACATGTGTAGAGGACAGAGATACAACTGGTCTTTCAATCA 181
QY	2955	CTGTCCAGCTTGGCAATGCAACGGCCACAGTAATGCATCAATCAGAGCATCTGCAGAA 3014
DB	182	CTGTCCAGCTTGGCAATGCAACGGCCACAGTAATGCATCAATCAGAGCATCTGCAGAA 241
QY	3015	GTGTGGAACCTGACCAAGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 3074
DB	242	GTGTGGAACCTGACCAAGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGA 301
QY	3075	TCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCATGGCAGCCGCTCTCTGTGCAA 3134
DB	302	TCCACCAATGGAGGGAAATGTCAGCCATGCAAGTGCATGGCAGCCGCTCTCTGTGCAA 361
QY	3135	CACCAACACGGGCAAGTGTCTTCGCACCAAGGGCGTCAAGGGGGAAGAGTCCAGCT 3194
DB	362	CACCAACACGGGCAAGTGTCTTCGCACCAAGGGCGTCAAGGGGGAAGAGTCCAGCT 421
QY	3195	ATGTGAGGTAGAAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATTACTCT 3254
DB	422	ATGTGAGGTAGAAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATTACTCT 481
QY	3255	TCCTATTGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGC 3314
DB	482	TCCTATTGACTATCAGTTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGC 541
QY	3315	TATCAATTTTGGCTACTCTCTCAGCAACAAAACAGGATTTGCAATGTCATCAATGC 3374
DB	542	TATCAATTTTGGCTACTCTCTCAGCAACAAAACAGGATTTGCAATGTCATCAATGC 601
QY	3375	CTCCAGAAATTTCAAACCTCAACATCAGTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGC 3434
DB	602	CTCCAGAAATTTCAAACCTCAACATCAGTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGC 661
QY	3435	TGAGAGAAGATGCCTGTTGTTTCAAAAACCAACATTAAAGGAGTACAAAGATAGTTTCTC 3494
DB	662	TGAGAGAAGATGCCTGTTGTTTCAAAAACCAACATTAAAGGAGTACAAAGATAGTTTCTC 721
QY	3495	TAATGAGAAGTTGATTTTCGCAACCAACCAATATACATCTTTCTTTGTTATGTCAGTAA 3554
DB	722	TAATGAGAAGTTGATTTTCGCAACCAACCAATATACATCTTTCTTTGTTATGTCAGTAA 780
QY	3555	TTTTCACCTGG 3564
DB	781	TTTTCACCTGG 790
RESULT 7		
LOCUS	CB520374	
DEFINITION	UI-M-G10-cei-1-18-0-UI.r1 NIH BMAP_G10 Mus musculus cDNA clone	linear EST 09-JUL-2003
ACCESSION	CB520374	
VERSION	CB520374.1	GI:29353729
KEYWORDS		
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

RESULT 7	CB520374	CB520374	CB520374	UI-M-G10-cei.1-18-0-UI.r1 NIH BMAP_G10 Mus musculus cDNA clone	linear	EST 09-JUL-2003
LOCUS	CB520374					
DEFINITION						
				IMAGE:6840067 5', mRNA sequence.		
ACCESSION						
VERSION						
KEYWORDS						
SOURCE						
ORGANISM						



LOCUS BM783788 683 bp mRNA linear EST 05-MAR-2002  
DEFINITION K-EST0061794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',  
mRNA sequence.  
ACCESSION BM783788  
VERSION BM783788.1 GI:19132020  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 683)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 36 row: E column: 09  
High quality sequence stop: 683.  
Location/Qualifiers  
FEATURES  
source  
1..683  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S5SNU484-36-E09"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/clone\_lib="S5SNU484"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site: 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was decapped with tobacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dT-selected  
mRNA by priming with dr-tailed vector. The dr-tailed  
vector was adjusted to have about 60nt. The cDNA vector  
was circularized with E. coli DNA ligase after digestion  
of EcoRI which site is also included in vector. An RNA  
strand converted to a DNA strand by Okayama-Berg method.  
The obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation  
method."

ORIGIN  
Query Match 16.7%; Score 680.4; DB 12; Length 683;  
Best Local Similarity 99.7%; Pred. No. 8.7e-162;  
Matches 681; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 2290 CAGCTGGCAATATGCAAGTCACTCAGAGCATGTCACCTTAACCCCATGGTTC 2349  
Db 1 CAGCTGGCAATATGCAAGTCACTCAGAGCATGTCACCTTAACCCCATGGTTC 60  
Qy 2350 GGCCTTCGGAAGATCAATGTCTCTACTGGTGTGGGAAGATATGTCCTCCATTTACAAAT 2409  
Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGGTGTGGGAAGATATGTCCTCCATTTACAAAT 120  
Qy 2410 AGTTTACTACAGTGGATGCGGCTGTAGCCAGTATGCTGGATCTGTGGAAATTTATCA 2469  
Db 121 AGTTTACTACAGTGGATGCGGCTGTAGCCAGTATGCTGGATCTGTGGAAATTTATCA 180  
Qy 2470 GAACCCAGTACTCGGGGACTGAAGCTGCAACTGCAATCAACCACTCAATGGTAGTGTG 2529  
Db 181 GAACCCAGTACTCGGGGACTGAAGCTGCAACTGCAATCAACCACTCAATGGTAGTGTG 240  
Qy 2530 TGTGAAAGGCTCGAAACCAACAGTGTCTAAGAGTGGCGGACACCATGTGCTTGGAGCA 2589

Db 241 TGTGAAAGGCTCGAAACCAACAGTGTCTAAGAGTGGCGGACACCATGTGCTTGGAGCA 300  
Qy 2590 GCATGTGAGATGTCACCGAGGAGCTCTGAGTGTGATGTGTCAGCAACATGAAGCAG 2649  
Db 301 GCATGTGAGATGTCACCGAGGAGCTCTGAGTGTGATGTGTCAGCAACATGAAGCAG 360  
Qy 2650 TGTGTGGACTCCAATGCTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2709  
Db 361 TGTGTGGACTCCAATGCTATGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 420  
Qy 2710 ACGATGAGACCTGCCCTCCCTGAAAAATTTGTTGAGGCTACTGTACTGTAGTGTGTTG 2769  
Db 421 ACGATGAGACCTGCCCTCCCTGAAAAATTTGTTGAGGCTACTGTACTGTAGTGTGTTG 480  
Qy 2770 GAGCAACAGGCTGTGGCTGTGCTGTACTGATCCAGCAATATCTGCAAGGGAATGCATA 2829  
Db 481 GAGCAACAGGCTGTGGCTGTGCTGTACTGATCCAGCAATATCTGCAAGGGAATGCATA 540  
Qy 2830 GAGGTTCTCTATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTCAGAGAAATTTCTAT 2889  
Db 541 GAGGTTCTCTATAAAGGACCAAGTGAAGATGCTTCGCAAGCCCTCAGAGAAATTTCTAT 600  
Qy 2890 CCAGAGCCCTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAACTGCTCTTC 2949  
Db 601 CCAGAGCCCTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAACTGCTCTTC 660  
Qy 2950 ATTCACTGTCCAGCTTGCCAAATG 2972  
Db 661 ATTCACTGTCCAGCTTGCCAAATG 683  
RESULT\_10  
BM783258  
LOCUS BM783258 680 bp mRNA linear EST 05-MAR-2002  
DEFINITION K-EST0061157 S5SNU484 Homo sapiens cDNA clone S5SNU484-31-H08 5',  
mRNA sequence.  
ACCESSION BM783258  
VERSION BM783258.1 GI:19131490  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 680)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 31 row: H column: 08  
High quality sequence stop: 680.  
Location/Qualifiers  
FEATURES  
source  
1..680  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S5SNU484-31-H08"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/clone\_lib="S5SNU484"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site: 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was decapped with tobacco  
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
including EcoRI site by treatment of T4 RNA ligase. The  
first strand cDNA was synthesized from oligo dT-selected  
mRNA by priming with dr-tailed vector. The dr-tailed  
vector was adjusted to have about 60nt. The cDNA vector  
was circularized with E. coli DNA ligase after digestion  
of EcoRI which site is also included in vector. An RNA  
strand converted to a DNA strand by Okayama-Berg method.  
The obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation  
method."

acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F<sup>+</sup> by electroporation method."

## ORIGIN

Query Match 16.7%; Score 680; DB 12; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.1e-161;  
Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2290 CAGCTCGGAATATGAGTCATCTCAGAGCATGCTCAAGCTACCTTAACCCCATGGGTC 2349  
DB 1 CAGCTCGGAATATGAGTCATCTCAGAGCATGCTCAAGCTACCTTAACCCCATGGGTC 60

QY 2350 GGCTTCGGGAATCAATGTGCTCTACTGCTGGGAAGATATGTCCTCAATTTCAAAAT 2409  
DB 61 GGCTTCGGGAATCAATGTGCTCTACTGCTGGGAAGATATGTCCTCAATTTCAAAAT 120

QY 2410 AGTTTACTACAGTGGATGCGCTGAGCCAGTGTCTGGATTTCTGTGGATTTTATCA 2469  
DB 121 AGTTTACTACAGTGGATGCGCTGAGCCAGTGTCTGGATTTCTGTGGATTTTATCA 180

QY 2470 GAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC 2529  
DB 181 GAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC 240

QY 2530 TGTGAAGGCTGCAACCAACAGTGTCTAGCAGTGGCGACCACTGTCCTTGAGGACA 2589  
DB 241 TGTGAAGGCTGCAACCAACAGTGTCTAGCAGTGGCGACCACTGTCCTTGAGGACA 300

QY 2590 GCATGTGGAGATTGCACCGGCGAGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAG 2649  
DB 301 GCATGTGGAGATTGCACCGGCGAGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAG 360

QY 2650 TGTGTGAGCTCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTGAAGTGGTAT 2709  
DB 361 TGTGTGAGCTCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTGAAGTGGTAT 420

QY 2710 ACCATGAGCAGCTGCCCCCTGAAATGTTTCAAGGCTACTGTACCTGTAGTCTGCTTG 2769  
DB 421 ACCATGAGCAGCTGCCCCCTGAAATGTTTCAAGGCTACTGTACCTGTAGTCTGCTTG 480

QY 2770 GAGCAACCGAGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGAAATGCATA 2829  
DB 481 GAGCAACCGAGCTGTGGCTGTACTGTATCCAGCAATACTGGCAAGGAAATGCATA 540

QY 2830 GAGGTTCTTATAAGGACCAAGTGAAGATGCTTCCGAGCCCTACAGGAAATTTCTAT 2889  
DB 541 GAGGTTCTTATAAGGACCAAGTGAAGATGCTTCCGAGCCCTACAGGAAATTTCTAT 600

QY 2890 CCAGAGCCCTGCTCAATTCAGCATGTCTAGAGGACAGCAGATACACTGGTCTTTC 2949  
DB 601 CCAGAGCCCTGCTCAATTCAGCATGTCTAGAGGACAGCAGATACACTGGTCTTTC 660

QY 2950 ATTCACGTGTCAGCTTGCCA 2969  
DB 661 ATTCACGTGTCAGCTTGCCA 680

RESULT 11  
BM783773  
LOCUS  
DEFINITION K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5', mRNA sequence.  
ACCESSION BM783773  
VERSION BM783773.1 GI:19132005

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi; Eukaryota; Metazoa; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 680)  
AUTHORS Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
COMMENT Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 36 row: D column: 02  
High quality sequence stop: 680.

## FEATURES

source  
1..680  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="S5SNU484-36-D02"  
/sex="M"  
/tissue\_type="Stomach"  
/cell\_type="Epithelial"  
/cell\_line="SNU-484"  
/lab\_host="Top10F"  
/clone\_lib="S5SNU484"  
/note="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-selected mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F<sup>+</sup> by electroporation method."

## ORIGIN

Query Match 16.7%; Score 680; DB 12; Length 680;  
Best Local Similarity 100.0%; Pred. No. 1.1e-161;  
Matches 680; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2290 CAGCTCGGAATATGAGTCATCTCAGAGCATGCTCAAGCTACCTTAACCCCATGGGTC 2349  
DB 1 CAGCTCGGAATATGAGTCATCTCAGAGCATGCTCAAGCTACCTTAACCCCATGGGTC 60

QY 2350 GGCTTCGGGAATCAATGTGCTCTACTGCTGGGAAGATATGTCCTCAATTTCAAAAT 2409  
DB 61 GGCTTCGGGAATCAATGTGCTCTACTGCTGGGAAGATATGTCCTCAATTTCAAAAT 120

QY 2410 AGTTTACTACAGTGGATGCGCTGAGCCAGTGTCTGGATTTCTGTGGATTTTATCA 2469  
DB 121 AGTTTACTACAGTGGATGCGCTGAGCCAGTGTCTGGATTTCTGTGGATTTTATCA 180

QY 2470 GAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC 2529  
DB 181 GAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC 240

QY 2530 TGTGAAGGCTGCAACCAACAGTGTCTAGCAGTGGCGACCACTGTCCTTGAGGACA 2589  
DB 241 TGTGAAGGCTGCAACCAACAGTGTCTAGCAGTGGCGACCACTGTCCTTGAGGACA 300

QY 2590 GCATGTGGAGATTGCACCGGCGAGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAG 2649



Db 301 GCATGTGAGATTGCACAGCGGAGCTCTCAGTGTGATGTGGTGCAGCAACATGAAGCAG 360  
 QY 2650 TGTGTGACTCCATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 2709  
 Db 361 TGTGTGACTCCATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 420  
 QY 2710 ACATGATGAGCACTGCCCTCCCTGAAATTTGTTCCAGTCTACTGTACCTGTAGTCTGCTTG 2769  
 Db 421 ACATGATGAGCACTGCCCTCCCTGAAATTTGTTCCAGTCTACTGTACCTGTAGTCTGCTTG 480  
 QY 2770 GAGCAACAGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCATA 2829  
 Db 481 GAGCAACAGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCATA 540  
 QY 2830 GAGGTTCTCTATAAGGACAGTGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTAT 2889  
 Db 541 GAGGTTCTCTATAAGGACAGTGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTAT 600  
 QY 2890 CCACAGCCCTGTCTCAATCCAGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTC 2949  
 Db 601 CCACAGCCCTGTCTCAATCCAGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTC 660  
 QY 2950 ATTCACTGTCCAGCTTGCCA 2969  
 Db 661 ATTCACTGTCCAGCTTGCCA 680

## RESULT 12

BM783718  
 LOCUS K-EST0061702 S5SNU484 Homo sapiens cDNA clone S5SNU484-34-F01 5',  
 DEFINITION mRNA sequence.

ACCESSION BM783718

VERSION BM783718.1 GI:19131950

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 677)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

AUTHORS Kim,N.S., Hahn,Y., Oh,J.H.Y., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
 Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
 Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsung@mail.kribb.re.kr

Plate: 34 row: F column: 01

High quality sequence stop: 677.

Location/Qualifiers

1..677

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S5SNU484-34-F01"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Epithelial"

/lab\_host="SNU-484"

/lab\_host="Top10P"

/clone\_lib="S5SNU484"

/note="Organ: Stomach; Vector: pTZ18RPI; Site: 1: EcoRI;  
 Site: 2: NotI; The poly (A) + RNA was decapped with tabacco  
 acid pyrophosphatase (TAP) and ligated with DNA-RNA linker  
 including EcoRI site by treatment of T4 RNA ligase. The  
 first strand cDNA was synthesized from oligo dT-selected  
 mRNA by priming with dT-tailed vector. The dT-tailed  
 vector was adjusted to have about 60nt. The cDNA vector

was circularized with E. coli DNA ligase after digestion  
 of EcoRI which site is also included in vector. An RNA  
 strand converted to a DNA strand by Okayama-Berg method.  
 The obtained cDNA vectors were used for transformation of  
 competent cells E. coli Top10P by electroporation  
 method."

## ORIGIN

Query Match 16.6%; Score 677; DB 12; Length 677;  
 Best Local Similarity 100.0%; Pred. No. 6.4e-161;  
 Matches 677; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2290 CAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAGCTCACTTAACCCCATGGGTC 2349  
 Db 1 CAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAGCTCACTTAACCCCATGGGTC 60  
 QY 2350 GGCCTTCGGAAGATCAATGTCTCTACTGTGTCGGGAAGATATGCCCATTTACCAAT 2409  
 Db 61 GGCCTTCGGAAGATCAATGTCTCTACTGTGTCGGGAAGATATGCCCATTTACCAAT 120  
 QY 2410 AGTTTACTACAGTGGATGCCCTCTCAGCCCAAGTGTGATTTCTGTGGAATTTATCA 2469  
 Db 121 AGTTTACTACAGTGGATGCCCTCTCAGCCCAAGTGTGATTTCTGTGGAATTTATCA 180  
 QY 2470 GAACCCAGTACTCGGGAGTGAAGCTGCAACCTGCATCAACCCCAAGTGTGATTTATCA 2529  
 Db 181 GAACCCAGTACTCGGGAGTGAAGCTGCAACCTGCATCAACCCCAAGTGTGATTTATCA 240  
 QY 2530 TGTGAAAGCCCTGCAAAACCAAGTGTAAAGCAGTGGCGGACACCATGTGCCCTCAGGACA 2589  
 Db 241 TGTGAAAGCCCTGCAAAACCAAGTGTAAAGCAGTGGCGGACACCATGTGCCCTCAGGACA 300  
 QY 2590 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 2649  
 Db 301 GCATGTGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAG 360  
 QY 2650 TGTGTGGAATCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 2709  
 Db 361 TGTGTGGAATCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 420  
 QY 2710 ACATGATGAGCACTGCCCTCCCTGAAATTTGTTCCAGTCTACTGTACCTGTAGTCTGCTTG 2769  
 Db 421 ACATGATGAGCACTGCCCTCCCTGAAATTTGTTCCAGTCTACTGTACCTGTAGTCTGCTTG 480  
 QY 2770 GAGCAACAGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCATA 2829  
 Db 481 GAGCAACAGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGGAATGCATA 540  
 QY 2830 GAGGTTCTCTATAAGGACAGTGAAGATGCTTTCGCAAGCCCTCAGGAAATTTCTAT 2889  
 Db 541 GAGGTTCTCTATAAGGACAGTGAAGATGCTTTCGCAAGCCCTCAGGAAATTTCTAT 600  
 QY 2890 CCACAGCCCTGTCTCAATCCAGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTC 2949  
 Db 601 CCACAGCCCTGTCTCAATCCAGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTC 660  
 QY 2950 ATTCACTGTCCAGCTTG 2966  
 Db 661 ATTCACTGTCCAGCTTG 677

## RESULT 13

BM783821

LOCUS K-EST0061830 S5SNU484

DEFINITION Homo sapiens cDNA clone S5SNU484-36-H09 5',

mRNA sequence.

ACCESSION BM783821

VERSION BM783821.1 GI:19132053

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.



method."

## ORIGIN

Query Match	16.5%;	Score 673;	DB 12;	Length 673;
Best Local Similarity	100.0%;	Prod. No. 6.7e-160;		
Matches 673;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	2290	CAGCTGCGAATAATGCAGTCTATCTCAGAGCATGTCCAAAGCTCACCTTTAAACCCATGGGTC	2349	
DB	1	CAGCTGCGAATAATGCAGTCTATCTCAGAGCATGTCCAAAGCTCACCTTTAAACCCATGGGTC	60	
QY	2350	GGCCTTTCGGAAGATCAATGTGTGCTCTA CTGTGTCTGGGAAGATATGTCCCATTTACAAT	2409	
DB	61	GGCCTTTCGGAAGATCAATGTGTGCTCTA CTGTGTCTGGGAAGATATGTCCCATTTACAAT	120	
QY	2410	AGTTTACTACAGTGGATGCCGTCTGAGCCACAGTGAATGCTCGAATTCCTGCAATTTTATCA	2469	
DB	121	AGTTTACTACAGTGGATGCCGTCTGAGCCACAGTGAATGCTCGAATTCCTGCAATTTTATCA	180	
QY	2470	GAACCCAGTACTCGGGACTGAAGGGCTGGAACCTGTGCATCAACCGACCTCAATGGTAGGTGTC	2529	
DB	181	GAACCCAGTACTCGGGACTGAAGGGCTGGAACCTGTGCATCAACCGACCTCAATGGTAGGTGTC	240	
QY	2530	TGTGAAAGGCGCTGCAAAACCAACAGTGTCTTAAGCAGTGGCGGACACCATGTGCTTTGAGGACA	2589	
DB	241	TGTGAAAGGCGCTGCAAAACCAACAGTGTCTTAAGCAGTGGCGGACACCATGTGCTTTGAGGACA	300	
QY	2590	GCATGTGGAGATTGCACACGGGAGCTCTGAGTGCATGTGGTGACAGCAACATGAAGCAG	2649	
DB	301	GCATGTGGAGATTGCACACGGGAGCTCTGAGTGCATGTGGTGACAGCAACATGAAGCAG	360	
QY	2650	TGTGTGACATCCAAATGCTATGTGGCTCCTCTCCCTTTTGGCCAGTGATGGAATGGTAT	2709	
DB	361	TGTGTGACATCCAAATGCTATGTGGCTCCTCTCCCTTTTGGCCAGTGATGGAATGGTAT	420	
QY	2710	ACGATGAGACCTGCCCCCTCGAAAATGTTTCAGGGTACTGTACCTGTACTCATTTGCTTG	2769	
DB	421	ACGATGAGACCTGCCCCCTCGAAAATGTTTCAGGGTACTGTACCTGTACTCATTTGCTTG	480	
QY	2770	GAGCAACACGGCTGTGGCTGTGTACTGATCCAGAGCAATCTGGCAAGGGAATGCGATA	2829	
DB	481	GAGCAACACGGCTGTGGCTGTGTACTGATCCAGAGCAATCTGGCAAGGGAATGCGATA	540	
QY	2830	GAGGGTTCTTATAAAGGACCAAGTGAAGATGCCCTTCGCAAGCCCTCACAGAAATTTCTAT	2889	
DB	541	GAGGGTTCTTATAAAGGACCAAGTGAAGATGCCCTTCGCAAGCCCTCACAGAAATTTCTAT	600	
QY	2890	CCACAGCCCTGTCTCAATTCGACATGTGTCATAGAGGACAGAGATACACTGGTCTTTC	2949	
DB	601	CCACAGCCCTGTCTCAATTCGACATGTGTCATAGAGGACAGAGATACACTGGTCTTTC	660	
QY	2950	ATTCACTGTCCAG 2962		
DB	661	ATTCACTGTCCAG 673		

DEPT. T 15

CD644584	792 bp	mRNA	linear	EST 17-JUN-2003
AGENCOURT_14539338	NIA Human H1 Embryonic Stem Cell	cDNA Library		
(Long)	Homo sapiens cDNA clone IMAGE:30421895	5', mRNA sequence.		
CD644584				
CD644584.1	GI:31816608			
EST.				
KEYWORDS				
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
AUTHORS	1. (bases 1 to 792)			
TITLE	NIH-MGC <a href="http://mgc.nci.nih.gov/">http://mgc.nci.nih.gov/</a> .			
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)			
COMMENT	Unpublished (1999)			
	Contact: Daniela S. Gerhard Ph.D.			



***This Page Blank (uspto)***



CC The present sequence is a cDNA encoding human membrane attractin-1.  
CC Membrane attractin differs from soluble attractin in having a  
CC transmembrane domain and a cytoplasmic domain at the C-terminal. This  
CC sequence can be used to enhance immune response in immunosuppressed  
CC patients such as those undergoing chemo- and radio-therapy treatment for  
CC cancer or those suffering from common variable immunodeficiency syndrome.  
CC The proteins may also be used to screen modulators (agonists and  
CC antagonists) of immune response which may also be used to regulate immune  
CC reactions. Attractin antibodies can be used to inhibit immune response in  
CC transplant recipients or patients afflicted with autoimmune disease  
XX  
SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;  
Query Match 100.0%; Score 4068; DB 3; Length 4068;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 4068; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGTTGCGCGAGCGCGGCAACTGAGGCAAGCTGAGGAGGAGGACGGCGGCGAGCGCA 60  
DB |||||  
QY 61 GCGCTCGCGGCGAGGCGGCGCGCACTGCTGCAAGCGGCTGCTGCAACCTCTGGC 120  
DB |||||  
QY 61 GCGCTCGCGGCGAGGCGGCGCGCACTGCTGCAAGCGGCTGCTGCAACCTCTGGC 120  
DB |||||  
QY 121 ACCGGCCAGTGCCTCTGCGCGCGCGGCTGGGTGGGCGAGCAATGCCAGCACTGCCGGGCG 180  
DB |||||  
QY 121 ACCGGCCAGTGCCTCTGCGCGCGCGGCTGGGTGGGCGAGCAATGCCAGCACTGCCGGGCG 180  
DB |||||  
QY 181 CGCTTCAGACTAAGTGAATCTTGGTGTGCAAGAGTGGACCTGGAATTAATAATAC 240  
DB |||||  
QY 181 CGCTTCAGACTAAGTGAATCTTGGTGTGCAAGAGTGGACCTGGAATTAATAATAC 240  
DB |||||  
QY 241 AATAAGAGTGCAGTGGCTCATTTGAAGGACAGCAAAATAGATAATAGACTTCGTTTC 300  
DB |||||  
QY 241 AATAAGAGTGCAGTGGCTCATTTGAAGGACAGCAAAATAGATAATAGACTTCGTTTC 300  
DB |||||  
QY 301 AATCATTTTGTCTACAGAGTGTAGTGGACCAATTTATGTTTATGATGGGAGCTCAATT 360  
DB |||||  
QY 301 AATCATTTTGTCTACAGAGTGTAGTGGACCAATTTATGTTTATGATGGGAGCTCAATT 360  
DB |||||  
QY 361 TATGACCGCTAGTGTGCTGATTTAGTGGCTCATTTGCTCAGAGAGATGCCAATGAG 420  
DB |||||  
QY 361 TATGACCGCTAGTGTGCTGATTTAGTGGCTCATTTGCTCAGAGAGATGCCAATGAG 420  
DB |||||  
QY 421 ACTGTCCCTCAGGTGTGTCACATCAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB |||||  
QY 421 ACTGTCCCTCAGGTGTGTCACATCAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCT 480  
DB |||||  
QY 481 GCTTATTAATTTGCTGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540  
DB |||||  
QY 481 GCTTATTAATTTGCTGATTTATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 540  
DB |||||  
QY 541 TCAGGCGGAGGAGTGTAAAGTCACTAATAGCAGGATCTCTGCAATGTGAATGTTCT 600  
DB |||||  
QY 541 TCAGGCGGAGGAGTGTAAAGTCACTAATAGCAGGATCTCTGCAATGTGAATGTTCT 600  
DB |||||  
QY 601 GAAACCTGGAAGGTGAAGCATGTGACATTCCTCAGTACAGCACTGCTGGTTCCT 660  
DB |||||  
QY 601 GAAACCTGGAAGGTGAAGCATGTGACATTCCTCAGTACAGCACTGCTGGTTCCT 660  
DB |||||  
QY 661 CATCGAGGCTCTGCAATTAAGTGTGTCAGAGGATGCTCTGCTGCTCTCAGACTGGCAG 720  
DB |||||  
QY 661 CATCGAGGCTCTGCAATTAAGTGTGTCAGAGGATGCTCTGCTGCTCTCAGACTGGCAG 720  
DB |||||  
QY 721 GGTCTCGGATGTCAGTTCCTGACAGCTAACAGCTCAATTTGGACTCGAGAGGAATAT 780  
DB |||||  
QY 721 GGTCTCGGATGTCAGTTCCTGACAGCTAACAGCTCAATTTGGACTCGAGAGGAATAT 780  
DB |||||  
QY 781 TCTAACTTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGCTCAATGGAACATTTATGCG 840  
DB |||||  
QY 781 TCTAACTTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGCTCAATGGAACATTTATGCG 840  
DB |||||

841 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGTTCTTAGCGTATGACCTT 900  
DB |||||  
841 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGTTCTTAGCGTATGACCTT 900  
DB |||||  
901 GCTTCTAGGAGTGTGCTTCCACTAAACCGTCTGCTGAAACATGTTGGTGTAGATATGTT 960  
DB |||||  
901 GCTTCTAGGAGTGTGCTTCCACTAAACCGTCTGCTGAAACATGTTGGTGTAGATATGTT 960  
DB |||||  
961 CATCTTTGGCATTTATCAAGGATATAAATTTACATGATATGAGGAAAAATTTGATCAACT 1020  
DB |||||  
961 CATCTTTGGCATTTATCAAGGATATAAATTTACATGATATGAGGAAAAATTTGATCAACT 1020  
DB |||||  
1021 GGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGATCATGGTGTGTTG 1080  
DB |||||  
1021 GGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGATCATGGTGTGTTG 1080  
DB |||||  
1081 ACCCTTAAGGCAAGGAGGAGTATGAGTGTGCTGGGCACTCTGCACACATTTGTACACTG 1140  
DB |||||  
1081 ACCCTTAAGGCAAGGAGGAGTATGAGTGTGCTGGGCACTCTGCACACATTTGTACACTG 1140  
DB |||||  
1141 AAGAAATGCGCGAGTGTCTGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATA 1200  
DB |||||  
1141 AAGAAATGCGCGAGTGTCTGCTGCTCATCTTTGGTCACTGCCCTCTCTATGGATATA 1200  
DB |||||  
1201 AGCAATGTGAGGAAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGT 1260  
DB |||||  
1201 AGCAATGTGAGGAAATATGATTTGGATAAGAACACATGGAGTATATTACACACCCAGGT 1260  
DB |||||  
1261 GCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTCTTACGACCATAGGACACGGGCCCTA 1320  
DB |||||  
1261 GCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTCTTACGACCATAGGACACGGGCCCTA 1320  
DB |||||  
1321 TACGTTCAATGTGGCTACAGGCTTTTCAAGTGCCTTCAAGTACCGGCTTGCAGATGATCTC 1380  
DB |||||  
1321 TACGTTCAATGTGGCTACAGGCTTTTCAAGTGCCTTCAAGTACCGGCTTGCAGATGATCTC 1380  
DB |||||  
1381 TACGATATGATGTGGATACCCAGATGTGACAGTCTTAAAGGACAGCCATTTTCCGT 1440  
DB |||||  
1381 TACGATATGATGTGGATACCCAGATGTGACAGTCTTAAAGGACAGCCATTTTCCGT 1440  
DB |||||  
1441 TACTTGCAACACAGCTGTGATAGTGGAAACCAATGCTGTTGGTGGGGAACACACAC 1500  
DB |||||  
1441 TACTTGCAACACAGCTGTGATAGTGGAAACCAATGCTGTTGGTGGGGAACACACAC 1500  
DB |||||  
1501 AATGACACATCTATGAGCCATGCGCCAAATGCTTCTTCAAGATTTTCAATGCTATGAC 1560  
DB |||||  
1501 AATGACACATCTATGAGCCATGCGCCAAATGCTTCTTCAAGATTTTCAATGCTATGAC 1560  
DB |||||  
1561 ATTGCTGTGACCGCTGCTGCTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGA 1620  
DB |||||  
1561 ATTGCTGTGACCGCTGCTGCTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGA 1620  
DB |||||  
1621 TTTGGCCATTACAGAGTCTTACAAACAGCAGCATGATGTGTTGGTGGTTCATAGT 1680  
DB |||||  
1621 TTTGGCCATTACAGAGTCTTACAAACAGCAGCATGATGTGTTGGTGGTTCATAGT 1680  
DB |||||  
1681 CTCCTCTCTCAGCAGCATCTCTGCTTATTCACCTCGGAAACAGTGTGATGCGCATCGAGTGA 1740  
DB |||||  
1681 CTCCTCTCTCAGCAGCATCTCTGCTTATTCACCTCGGAAACAGTGTGATGCGCATCGAGTGA 1740  
DB |||||  
1741 GCGGCTTGTAGCAGGAGGACCTGCTGATTTGCTGTGTGTGGACACAGGCTGCTCTCAG 1800  
DB |||||  
1741 GCGGCTTGTAGCAGGAGGACCTGCTGATTTGCTGTGTGTGGACACAGGCTGCTCTCAG 1800  
DB |||||  
1801 TGTATCTGCTGGCGCTGGCAATGATGAACAAGAAAGTTTAAATCAGATGTTTT 1860  
DB |||||  
1801 TGTATCTGCTGGCGCTGGCAATGATGAACAAGAAAGTTTAAATCAGATGTTTT 1860  
DB |||||  
1861 TCCAAAGAACTCTTGTGACCAATGACAGTGTGACGACACAGATTTGTACAGCTGCACA 1920  
DB |||||  
1861 TCCAAAGAACTCTTGTGACCAATGACAGTGTGACGACACAGATTTGTACAGCTGCACA 1920  
DB |||||  
1921 GCCAACCAATGACTGCCACTGCTGTCATGACCAATGCTGTGCTCCCGAGGAACACAGCTGC 1980  
DB |||||



Db 1921 GCCAACACCAATGACTGCCACTGGTGAATGACCAATGGTGTCCCGAGGAACACAGCTGC 1980  
Qy 1981 TCAGAAAGGCAGATCTCCATTTTATAGTATGAAATGGCCCAAGATACACCTTATGTAC 2040  
Db 1981 TCAGAAAGGCAGATCTCCATTTTATAGTATGAAATGGCCCAAGATACACCTTATGTAC 2040  
Qy 2041 TACTGTAAACAAGAACACAGCTGCAGAGCTGTGGCCCTGGACCAAGAACTGCCAGTGGAG 2100  
Db 2041 TACTGTAAACAAGAACACAGCTGCAGAGCTGTGGCCCTGGACCAAGAACTGCCAGTGGAG 2100  
Qy 2101 CCCCGGAATCAGAGTGCATTCGCCCTGCCGGAATATCTGTGGCAATGGCTGGCATTTG 2160  
Db 2101 CCCCGGAATCAGAGTGCATTCGCCCTGCCGGAATATCTGTGGCAATGGCTGGCATTTG 2160  
Qy 2161 GTTGGAAATCTCATGTTTGAATTTACTGTCGAAGGAAATTTATGACAAATGCTAAATTTG 2220  
Db 2161 GTTGGAAATCTCATGTTTGAATTTACTGTCGAAGGAAATTTATGACAAATGCTAAATTTG 2220  
Qy 2221 TTCTGTAGGAACACAAATGCCCTTTTGGCTTCTCTTACAAACCCAGAGGAGTAGAATTT 2280  
Db 2221 TTCTGTAGGAACACAAATGCCCTTTTGGCTTCTCTTACAAACCCAGAGGAGTAGAATTT 2280  
Qy 2281 GTCCCTTAAGCAGCTGCGAATATGCAAGTCTCTCAGAGCATGTCCAAAGCTCACCTTAACC 2340  
Db 2281 GTCCCTTAAGCAGCTGCGAATATGCAAGTCTCTCAGAGCATGTCCAAAGCTCACCTTAACC 2340  
Qy 2341 CCATGGTGGCCCTTCGGAAGATCAATGTGTCTACTGTGTGGGAAGATATGTCCTCCA 2400  
Db 2341 CCATGGTGGCCCTTCGGAAGATCAATGTGTCTACTGTGTGGGAAGATATGTCCTCCA 2400  
Qy 2401 TTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCAGTGTGCTGATTCGTGGA 2460  
Db 2401 TTTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCAGTGTGCTGATTCGTGGA 2460  
Qy 2461 ATTTTATCAGAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCACTCAACCACTCAAT 2520  
Db 2461 ATTTTATCAGAACCCAGTACTCGGGACTGAAGGCTGCAACCTGCACTCAACCACTCAAT 2520  
Qy 2521 GGTAGTGTCTGTAAAGGCTGCAACCAAGTGTAAAGGCTGCGGACCAAGTGTGCC 2580  
Db 2521 GGTAGTGTCTGTAAAGGCTGCAACCAAGTGTAAAGGCTGCGGACCAAGTGTGCC 2580  
Qy 2581 TTGAGCAGCAGTGTGAGATTCACACAGCGGAGCTCTGAGTGCATGTGTGCGAGCAAC 2640  
Db 2581 TTGAGCAGCAGTGTGAGATTCACACAGCGGAGCTCTGAGTGCATGTGTGCGAGCAAC 2640  
Qy 2641 ATGAAGCAGTGTGAGCTCAATGTGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATG 2700  
Db 2641 ATGAAGCAGTGTGAGCTCAATGTGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATG 2700  
Qy 2701 GAATGGTATACGATGAGCACTGCCCTGCAATTTGTTTCAAGGCTACTGTACCTGTAGT 2760  
Db 2701 GAATGGTATACGATGAGCACTGCCCTGCAATTTGTTTCAAGGCTACTGTACCTGTAGT 2760  
Qy 2761 CATTTGTTGAGCAACAGGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGG 2820  
Db 2761 CATTTGTTGAGCAACAGGCTGTGGCTGTGTACTGATCCCAAGCAATACTGGCAAGGG 2820  
Qy 2821 AATATGATAGAGGTTCTTATTAAGGACAGTGAAGTGTGCAAGCCCTTACAGGA 2880  
Db 2821 AATATGATAGAGGTTCTTATTAAGGACAGTGAAGTGTGCAAGCCCTTACAGGA 2880  
Qy 2881 AATTTCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGACAGCATACAAC 2940  
Db 2881 AATTTCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGACAGCATACAAC 2940  
Qy 2941 TGGTCTTTTCTATCTGTGTCAGGTTGCAATGCAACGGGCAAGTAAATGCAATCAATCAG 3000  
Db 2941 TGGTCTTTTCTATCTGTGTCAGGTTGCAATGCAACGGGCAAGTAAATGCAATCAATCAG 3000  
Qy 3001 AGCATCTGTGAGAGTGTGAGAACCTTGACACAGGCAAGCACTGCCAGACCTGCATCTCT 3060

RESULT 2  
AAZ91920

Db 3001 AGCATCTGTGAGAGTGTGAGAACCTTGACCAAGCAAGCACTGCGAGACCTGCATCTCT 3060  
Qy 3061 GGCCTTCTACGGTGTATCCCAATGAGGGGAAATGTGAGCCATCAAGTGCATATGGGCAC 3120  
Db 3061 GGCCTTCTACGGTGTATCCCAATGAGGGGAAATGTGAGCCATCAAGTGCATATGGGCAC 3120  
Qy 3121 GCGTCTCTGTGCAACACCAACACCGGCAAGTGTCTTGACCCACCAAGGGCGTCAAGGGG 3180  
Db 3121 GCGTCTCTGTGCAACACCAACACCGGCAAGTGTCTTGACCCACCAAGGGCGTCAAGGGG 3180  
Qy 3181 GACGAGTGCCAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACA 3240  
Db 3181 GACGAGTGCCAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACA 3240  
Qy 3241 TGTATTATATCTCTCTTATTAGTATCAGTTCACCTTTTATGTCTATCCAGGAAGATGAT 3300  
Db 3241 TGTATTATATCTCTCTTATTAGTATCAGTTCACCTTTTATGTCTATCCAGGAAGATGAT 3300  
Qy 3301 CGCTATTACAGCTATCAATTTTGTGGCTACTCTGACGAACAAACAGGGATTTGAC 3360  
Db 3301 CGCTATTACAGCTATCAATTTTGTGGCTACTCTGACGAACAAACAGGGATTTGAC 3360  
Qy 3361 ATGTTTCATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTCTCA 3420  
Db 3361 ATGTTTCATGCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTCTCA 3420  
Qy 3421 GCTGGAAACCCAGGCTGAGAGAGATGCTGTTTGTCTTCAAAAACCAACATTAAGAGTAC 3480  
Db 3421 GCTGGAAACCCAGGCTGAGAGAGATGCTGTTTGTCTTCAAAAACCAACATTAAGAGTAC 3480  
Qy 3481 AAAGATAGTTCTCTAATGAGAAATTTGATTTTCCGCAACCAACCAAAATATCACTTCTTT 3540  
Db 3481 AAAGATAGTTCTCTAATGAGAAATTTGATTTTCCGCAACCAACCAAAATATCACTTCTTT 3540  
Qy 3541 GTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGATTTGCTCTCTCAGCAGAC 3600  
Db 3541 GTTTATGTCAGTAATTTCACTGGCCCATCAAAATTCAGATTTGCTCTCTCAGCAGAC 3600  
Qy 3601 AATTTTATGAGACCTGTGTACAGTCTTGTGACTTCTTTCAGTGTCTCTCTCTTTGCTC 3660  
Db 3601 AATTTTATGAGACCTGTGTACAGTCTTGTGACTTCTTTCAGTGTCTCTCTCTTTGCTC 3660  
Qy 3661 CTGTGGCTGTGTGTTTGAAGATCAAAACAAAGTGTGTGGGCTCCAGACGTAGAGAG 3720  
Db 3661 CTGTGGCTGTGTGTTTGAAGATCAAAACAAAGTGTGTGGGCTCCAGACGTAGAGAG 3720  
Qy 3721 CAACCTTCTCAGAGATGCAACAGATGGCCAGCGTCCCTTTGCTCTGTAAATGTGCCC 3780  
Db 3721 CAACCTTCTCAGAGATGCAACAGATGGCCAGCGTCCCTTTGCTCTGTAAATGTGCCC 3780  
Qy 3781 TTGGAACACAGATGAGGAGCTCTGATCTTATTTGGGGGAGTATAAAGACTGTTCCTCAAA 3840  
Db 3781 TTGGAACACAGATGAGGAGCTCTGATCTTATTTGGGGGAGTATAAAGACTGTTCCTCAAA 3840  
Qy 3841 CCCATTTGACCTGGAGCCGTGTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGG 3900  
Db 3841 CCCATTTGACCTGGAGCCGTGTTTGGCAACAAAGCCGCTGTCCTCTCTGTGTTTGTGAGG 3900  
Qy 3901 CTCCTCCAGGCTGGGTGGCATCCCTCTCCGCGAGTCCAGTCTTGTGTGGCCAGC 3960  
Db 3901 CTCCTCCAGGCTGGGTGGCATCCCTCTCCGCGAGTCCAGTCTTGTGTGGCCAGC 3960  
Qy 3961 GCCCTGTGTGACATTTTCTCAGCAGATCCGATAGTGTACAAGGAGAAAGTCAAGGAGCCGTG 4020  
Db 3961 GCCCTGTGTGACATTTTCTCAGCAGATCCGATAGTGTACAAGGAGAAAGTCAAGGAGCCGTG 4020  
Qy 4021 AGAAACCGGAAGCAGCAGCCCCCTGCAAGCTGGACCTGCATCTGA 4068  
Db 4021 AGAAACCGGAAGCAGCAGCCCCCTGCAAGCTGGACCTGCATCTGA 4068

ID AA291920 standard; cDNA; 8590 BP.  
 XX AC AA291920;  
 XX DT 08-JUN-2000 (first entry)  
 XX DE Human mahogany protein coding sequence #2.  
 XX KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
 XX KW weight regulation; cell therapy; body weight disorder; cachexia;  
 XX KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 XX KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
 XX OS Homo sapiens.  
 XX XX  
 XX PN WO200005373-A2.  
 XX PD 03-FEB-2000.  
 XX PF 21-JUL-1999; 99WO-US016484.  
 XX PR 21-JUL-1998; 98US-0093630P.  
 XX PR 20-OCT-1998; 98US-0104978P.  
 XX PR 05-FEB-1999; 99US-00245041.  
 XX XX  
 XX PA (MILL-) MILLENIUM PHARM INC.  
 XX PI Moore K, Nagle DL;  
 XX DR WPI; 2000-195103/17.  
 XX DR P-PSDB; AAY81807.  
 XX PT New human and murine mahogany genes, useful, e.g. for diagnosis and  
 XX PT treatment of body weight disorders.  
 XX PS Claim 1; Fig 18a; 188pp; English.  
 XX XX  
 CC This sequence represents a human mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (ii) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (ii)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX SQ Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;  
 Query Match 98.0%; Score 3988.4; DB 3; Length 8590;  
 Best Local Similarity 99.0%; Pred. No. 0;  
 Matches 4013; Conservative 0; Mismatches 41; Indels 0; Gaps 0;  
 QY 15 GGCGGCAACTGAGCAAGGCTGAGGAGGAGGAGCGGGCGACCGGAGCGCTCGCGGGCAG 74  
 DB 234 GGAGCCGAGCGCCCTCGCGTGGCGCGGGTGTCCGGCTCGCGCGAGCGGCGCA 293  
 QY 75 GAGCGCGGCGCGCACTGTGTCAACCGCGGTGCGTCAACCCCTGGCAACCGCGCGAGTGCCT 134  
 DB 294 GGAATGTGACCGGCGGTGTGTCAACCGCGGTGCTGCAACCCCTGGCAACCGCGCGAGTGCCT 353  
 QY 135 CTGCCCCCGCGGTGGGTGGCGGAGCAATCCAGCACTGGGGGGCGGCTTCAGACTAAC 194  
 DB 354 CTGCCCCCGCGGTGGGTGGCGGAGCAATCCAGCACTGGGGGGCGGCTTCAGACTAAC 413

195 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATATATAATACAAAACGAGTGCAC 254  
 DB 414 TGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATATATAATACAAAACGAGTGCAC 473  
 QY 255 GTGGCTCATTCGAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTGAC 314  
 DB 474 GTGGCTCATTCGAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTGAC 533  
 QY 315 AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGT 374  
 DB 534 AGAGTGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTATGACCGCTAGT 593  
 QY 375 TGCTGCATTTAGTGGCCCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGT 434  
 DB 594 TGCTGCATTTAGTGGCCCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCTGAGGT 653  
 QY 435 TGTGGCCACATCAGGTTATGCTTGTGCTGCAATTTTTTTAGTGTGATGCTGCTTATATTTGAC 494  
 DB 654 TGTGGCCACATCAGGTTATGCTTGTGCTGCAATTTTTTTAGTGTGATGCTGCTTATATTTGAC 713  
 QY 495 TGGATTTAATATTACTTTACAGTTTGTATGTTGTCCTCAATTAACCTGCTCAGCCGAGGAGA 554  
 DB 714 TGGATTTAATATTACTTTACAGTTTGTATGTTGTCCTCAATTAACCTGCTCAGCCGAGGAGA 773  
 QY 555 GTGTAGATCAGTAATAGCAGCGATACCTGTTGAATGTGAATGTTCTGAAAACCTGGAAGG 614  
 DB 774 GTGTAGATCAGTAATAGCAGCGATACCTGTTGAATGTGAATGTTCTGAAAACCTGGAAGG 833  
 QY 615 TGAAGCATGTCACATTTCTCACTGTACAGACAACCTGTTGGTTTCTCATCGAGGCATCTG 674  
 DB 834 TGAAGCATGTCACATTTCTCACTGTACAGACAACCTGTTGGTTTCTCATCGAGGCATCTG 893  
 QY 675 CAATTCAGTGTGATGACAGGATGCTCTGCTTCTCAGACTGGGAGGCTCTGATGTTTC 734  
 DB 894 CAATTCAGTGTGATGACAGGATGCTCTGCTTCTCAGACTGGGAGGCTCTGATGTTTC 953  
 QY 735 AGTTCTGTACAGCTAACCACTCATTTTGGACTCGAGAGGAATATTTCTAACTTTAAAGCT 794  
 DB 954 AGTTCTGTACAGCTAACCACTCATTTTGGACTCGAGAGGAATATTTCTAACTTTAAAGCT 1013  
 QY 795 CCCAGAGCATCTCATAAAGCTGTGTCATGGAACAATTTATGTTGGTGTGTGGAGGATA 854  
 DB 1014 CCCAGAGCATCTCATAAAGCTGTGTCATGGAACAATTTATGTTGGTGTGTGGAGGATA 1073  
 QY 855 TATGTTCAACCACTCAGATATATACAGTGTCTAGCGTATGACCTGCTTCTAGGAGGTG 914  
 DB 1074 TATGTTCAACCACTCAGATATATACAGTGTGTAGCGTATGACCTGCTTCTAGGAGGTG 1133  
 QY 915 GCTTCCACTAAACCGTTCTGTGAAACAATGTTGTTGTTAGATATGTTCAATTTTGGCAAT 974  
 DB 1134 GCTTCCACTAAACCGTTCTGTGAAACAATGTTGTTGTTAGATATGTTCAATTTTGGCAAT 1193  
 QY 975 ATACAGATATAAATTTACATGTTAGAGGAAATTTGATTTCACTGGGATGTGACCA 1034  
 DB 1194 ATACAGATATAAATTTACATGTTAGAGGAAATTTGATTTCACTGGGATGTGACCA 1253  
 QY 1035 TGAGTTGAGAGTGTGTTTCACTTCAATTAATGAGTGTGTTGTTGACCCCTTAAGGCAAA 1094  
 DB 1254 TGAGTTGAGAGTGTGTTTCACTTCAATTAATGAGTGTGTTGTTGACCCCTTAAGGCAAA 1313  
 QY 1095 GGAGCAGTATGAGTGTGTTGGGCACTCTGTCACACATTTGTTTACACTGAAGATGCCCCAGT 1154  
 DB 1314 GGAGCAGTATGAGTGTGTTGGGCACTCTGTCACACATTTGTTTACACTGAAGATGCCCCAGT 1373  
 QY 1155 GGTCAATGCTGTGTCATCTTTGGTCACTGCGCTCTCTATGGATATATAAGCAATGTGCAGGA 1214  
 DB 1374 GGTCAATGCTGTGTCATCTTTGGTCACTGCGCTCTCTATGGATATATAAGCAATGTGCAGGA 1433  
 QY 1215 ATATGATTTGGATAAGAACACATCGAGTATATATACACCCGAGGTCCTTGTGCAAGG 1274  
 DB 1434 ATATGATTTGGATAAGAACACATCGAGTATATATACACCCGAGGTCCTTGTGCAAGG 1493

QY 1275 GGGTTACGGCCATAGCAGTGTGTTTACGACCATAGGACCAGGGCCCTATATAGCTTCATGGTGG 1334  
Db 1494 GGGTTACGGCCATAGCAGTGTGTTTACGACCATAGGACCAGGGCCCTATATAGCTTCATGGTGG 1553  
QY 1335 CTACAAGGCTTTTCAGTGGCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394  
Db 1554 CTACAAGGCTTTTCAGTGGCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1613  
QY 1395 GGATACCCAGATGTGACCATCTTTAAGGACAGCCGATTTTTCGGTACTTTCGACACAGC 1454  
Db 1614 GGATACCCAGATGTGACCATCTTTAAGGACAGCCGATTTTTCGGTACTTTCGACACAGC 1673  
QY 1455 TGTGATAGTGTGAGGAAACCATGCTGTGTTTGGGGGAAACACACAAATGACACATCTAT 1514  
Db 1674 TGTGATAGTGTGAGGAAACCATGCTGTGTTTGGGGGAAACACACAAATGACACATCTAT 1733  
QY 1515 GAGCCATGGCGCCAAATGTCTCTTTTCAGATTTTCATGGCTATGACATTTGCCGTGTGACCG 1574  
Db 1734 GAGCCATGGCGCCAAATGTCTCTTTTCAGATTTTCATGGCTATGACATTTGCCGTGTGACCG 1793  
QY 1575 CTGGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTTGGCCATTCAGC 1634  
Db 1794 CTGGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTTGGCCATTCAGC 1853  
QY 1635 AGTCTTACCAACAGCACCATGTATGTGTTTGGGTGTTTCAATAGTCTCTCTCCTCAGCGA 1694  
Db 1854 AGTCTTACCAACAGCACCATGTATGTGTTTGGGTGTTTCAATAGTCTCTCTCCTCAGCGA 1913  
QY 1695 CATCTGTGTTTCACTCTGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTAGC 1754  
Db 1914 CATCTGTGTTTCACTCTGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTAGC 1973  
QY 1755 AGCAGACCTGTGTTTGGGTGTTTGGAAACAGGCTGTCTCAGTGTATCTGCTGGGC 1814  
Db 1974 AGCAGACCTGTGTTTGGGTGTTTGGAAACAGGCTGTCTCAGTGTATCTGCTGGGC 2033  
QY 1815 GCTGGCAACTGTATGAAACAGAAAGTTTAAATCAGAAATGTTTTCAAAAGAACTCT 1874  
Db 2034 GCTGGCAACTGTATGAAACAGAAAGTTTAAATCAGAAATGTTTTCAAAAGAACTCT 2093  
QY 1875 TGACCATGACAGATGTGACGACACAGATTTGTATACAGTGCACAGCCAAACCCATGA 1934  
Db 2094 TGACCATGACAGATGTGACGACACAGATTTGTATACAGTGCACAGCCAAACCCATGA 2153  
QY 1935 CTGCCACTGTGCAATGTGCTCCAGGAACACAGCTGTCTCAGAAAGGCCAGAT 1994  
Db 2154 CTGCCACTGTGCAATGTGCTCCAGGAACACAGCTGTCTCAGAAAGGCCAGAT 2213  
QY 1995 CTCCTATTTTGTATGAAATGCCCCAAGGATAACCCATGTACTACTGTAAACAAGAA 2054  
Db 2214 CTCCTATTTTGTATGAAATGCCCCAAGGATAACCCATGTACTACTGTAAACAAGAA 2273  
QY 2055 GACCAGCTGACGAGCTGTGCTGACACAGAACTGCCAGTGGAGCCCGGAAATCAGGA 2114  
Db 2274 GACCAGCTGACGAGCTGTGCTGACACAGAACTGCCAGTGGAGCCCGGAAATCAGGA 2333  
QY 2115 GTGCATTTGCCCTCCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACCTCATG 2174  
Db 2334 GTGCATTTGCCCTCCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACCTCATG 2393  
QY 2175 TTTGAAATTTTACTACTGCCAAGGAAATTTATGCAATGCTAAATGTTCTGTAGGAACCA 2234  
Db 2394 TTTGAAATTTTACTACTGCCAAGGAAATTTATGCAATGCTAAATGTTCTGTAGGAACCA 2453  
QY 2235 CAATGGCCCTTTTGGCTTCTTTTACAAACCCAGAGAAAGGTAGAAATTTGCTCTTTAAGCAGCT 2294  
Db 2454 CAATGGCCCTTTTGGCTTCTTTTACAAACCCAGAGAAAGGTAGAAATTTGCTCTTTAAGCAGCT 2513  
QY 2295 GCGAAATTAATGCACTCATCTCAGAGCATGTCCAGCTCAGCTTAAACCCCATGGGTGGCCT 2354  
Db 2514 GCGAAATTAATGCACTCATCTCAGAGCATGTCCAGCTCAGCTTAAACCCCATGGGTGGCCT 2573  
QY 2355 TCGGAAGATCAATGTGTCTTACTGTGGTGTGGGAAAGATATGTCTCCCATTTTACAAATAGTTT 2414

Db 2574 TCGGAAGATCAATGTGTCTTACTGTGTCTGGGAAGATATGTCTCCCATTTTACAAATAGTTT 2633  
QY 2415 ACTACAGTGGATGCGCTCTGAGCCCAAGTGTCTGGATTTCTGTGGAAATTTTATCAGAACC 2474  
Db 2634 ACTACAGTGGATGCGCTCTGAGCCCAAGTGTCTGGATTTCTGTGGAAATTTTATCAGAACC 2693  
QY 2475 CAGTACTCGGGGACTGAAGGCTGCAACTGTGATCAAAACCACTCAATGGTAGTGTCTGTGA 2534  
Db 2694 CAGTACTCGGGGACTGAAGGCTGCAACTGTGATCAAAACCACTCAATGGTAGTGTCTGTGA 2753  
QY 2535 AAGGCTTCAAAACACAGTGTCTAAGCAGTGCAGACACCACTGTGCTTTGAGGACAGCATG 2594  
Db 2754 AAGGCTTCAAAACACAGTGTCTAAGCAGTGCAGACACCACTGTGCTTTGAGGACAGCATG 2813  
QY 2595 TGGAGATTTGACACAGCGGAGCTCTGAGTGCATGTGTGACGAAACAAAGAGCAGTGTGT 2654  
Db 2814 TGGAGATTTGACACAGCGGAGCTCTGAGTGCATGTGTGACGAAACAAAGAGCAGTGTGT 2873  
QY 2655 GGACTCCAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2714  
Db 2874 GGACTCCAATGCCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGAT 2933  
QY 2715 GAGCACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATTCCTTTGGAGCA 2774  
Db 2934 GAGCACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATTCCTTTGGAGCA 2993  
QY 2775 ACCAGGCTGTGGCTGTGTACTGATCCAGCAAACTACTGGCAAGGAAATGCAATAGAGGG 2834  
Db 2994 ACCAGGCTGTGGCTGTGTACTGATCCAGCAAACTACTGGCAAGGAAATGCAATAGAGGG 3053  
QY 2835 TTCTCTATAAAGGACACAGTGAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACA 2894  
Db 3054 TTCTCTATAAAGGACACAGTGAAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACA 3113  
QY 2895 GCGCTGTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCAATCA 2954  
Db 3114 GCGCTGTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTCAATCA 3173  
QY 2955 CTGTCCAGTGTGCAATGTCAACGGCCACAGTAAATGATCAATCAGAGCATCTCTGTAGAA 3014  
Db 3174 CTGTCCAGTGTGCAATGTCAACGGCCACAGTAAATGATCAATCAGAGCATCTCTGTAGAA 3233  
QY 3015 GTGTGAGAACCTGACCAAGCAGCAGTGTGAGACCTGCAATATCTGGCTTTCTACGGTGA 3074  
Db 3234 GTGTGAGAACCTGACCAAGCAGCAGTGTGAGACCTGCAATATCTGGCTTTCTACGGTGA 3293  
QY 3075 TCCACCAATGAGAGGAAATGTGAGCCATGCAAGTGCATTTGGGACCGCTCTCTGTGCAA 3134  
Db 3294 TCCACCAATGAGAGGAAATGTGAGCCATGCAAGTGCATTTGGGACCGCTCTCTGTGCAA 3353  
QY 3135 CACCAACACGGGCAAGTGTCTGTGACCAACCAAGGGGCTCAAGGGGACAGAGTGCACAGCT 3194  
Db 3354 CACCAACACGGGCAAGTGTCTGTGACCAACCAAGGGGCTCAAGGGGACAGAGTGCACAGCT 3413  
QY 3195 ATGTGAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTTACTCT 3254  
Db 3414 ATGTGAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTTACTCT 3473  
QY 3255 TCTTATTTGACTATCAGTTTCACTTTTGTCTATCCAGGAAGATGATCGCTATTTACACAGC 3314  
Db 3474 TCTTATTTGACTATCAGTTTCACTTTTGTCTATCCAGGAAGATGATCGCTATTTACACAGC 3533  
QY 3315 TATCAATTTTGTGCTACTCTCTGACGAACAAAACAGGGATTTGGACATGTTTCAATGTC 3374  
Db 3534 TATCAATTTTGTGCTACTCTCTGACGAACAAAACAGGGATTTGGACATGTTTCAATGTC 3593  
QY 3375 CTCGAAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434  
Db 3594 CTCGAAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653  
QY 3435 TGGGAAGATGCTGCTGTTTGTCTTCAAAACCAACATTAAGGAGTACAAAGATAGTTCTC 3494



810 ATGTGAATGTTCTGAAAACTGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGCAA 869  
648 CTGTGTTTTTCCTCATPCGAGGATCTGCAATTCAGATGATGTCAGAGATGCTCTCTGTT 707  
870 CTGTGTTTTTCCTCATPCGAGGATCTGCAATTCAGATGATGTCAGAGATGCTCTCTGTT 929  
708 CTCAGACTGGCAGGTCCTGGATGTTTCAGTTCCTGTPACAGCTAACCACTCAATTTGGAC 767  
930 CTCAGACTGGCAGGTCCTGGATGTTTCAGTTCCTGTPACAGCTAACCACTCAATTTGGAC 989  
768 TCAGAGGAATATTCCTAACTTAAAGCTCCCGAGAGCATCTATAAGCTGTGGTCAATGG 827  
990 TCAGAGGAATATTCCTAACTTAAAGCTCCCGAGAGCATCTATAAGCTGTGGTCAATGG 1049  
828 AATCATATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAAATGGTTCT 887  
1050 AATCATATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTATAAATGGTTCT 1109  
888 AGCGTATGACCTTCTCTAGGAGTGGCTTCACCTAAACCGTCTCTGTCGAACAATGTGGT 947  
1110 AGCGTATGACCTTCTCTAGGAGTGGCTTCACCTAAACCGTCTCTGTCGAACAATGTGGT 1169  
948 TGTAGATATGTCATCTCTTGGCATTATACAGGATATAAATTTACATGTATGGAGGAAA 1007  
1170 TGTAGATATGTCATCTCTTGGCATTATACAGGATATAAATTTACATGTATGGAGGAAA 1229  
1008 AATTGATTCAACTGGGAATGTGACCAATGATGATGAGATTTTTCACATTCATAATGATC 1067  
1230 AATTGATTCAACTGGGAATGTGACCAATGATGATGAGATTTTTCACATTCATAATGATC 1289  
1068 ATCGGTGTTGTGACCCCTTAAGCCAAAGGAGCATGATGAGTGGTGGGCATCTGTCACA 1127  
1290 ATCGGTGTTGTGACCCCTTAAGCCAAAGGAGCATGATGAGTGGTGGGCATCTGTCACA 1349  
1128 CATGTTTACACTGAAGAATGGCCGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1187  
1350 CATGTTTACACTGAAGAATGGCCGAGTGGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1409  
1188 CTATGATATATGAAGCAATGTGAGGAATATGATTTGGATAGAAGACACATGAGTATATT 1247  
1410 CTATGATATATGAAGCAATGTGAGGAATATGATTTGGATAGAAGACACATGAGTATATT 1469  
1248 ACACACCGAGGTGCCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTGTTTACACCATAG 1307  
1470 ACACACCGAGGTGCCCCTTGTGCAAGGGGTTACGGCCATAGCAGTGTGTTTACACCATAG 1529  
1308 GACCAGGGCCCTATACGTTTCATGGTGGCTACAAGGCTTTTCAGTCCCAATAGTACCGGCT 1367  
1530 GACCAGGGCCCTATACGTTTCATGGTGGCTACAAGGCTTTTCAGTCCCAATAGTACCGGCT 1589  
1368 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATCTTAAGGACAG 1427  
1590 TGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCATCTTAAGGACAG 1649  
1428 CCGATTTTTCGCTTACTTGCACACAGCTGATGATGAGTGGAGCAACATGCTGTTTGG 1487  
1650 CCGATTTTTCGCTTACTTGCACACAGCTGATGATGAGTGGAGCAACATGCTGTTTGG 1709  
1488 GGGAAAACACACAATGACATCTATGAGCCATGGCCCAATGCTTCTCTTCAGATTT 1547  
1710 GGGAAAACACACAATGACATCTATGAGCCATGGCCCAATGCTTCTCTTCAGATTT 1769  
1548 CATGGCCCTATGACATTCGCTGTGACCGCTGGTCACTGCTTCCAGACCTGATCTCCACCA 1607  
1770 CATGGCCCTATGACATTCGCTGTGACCGCTGGTCACTGCTTCCAGACCTGATCTCCACCA 1829  
1608 TGATGTCAACAGATTTGGCCATTCAGAGTCTTACACAACAGCAGCACCATGATGTGTCGG 1667  
1830 TGATGTCAACAGATTTGGCCATTCAGAGTCTTACACAACAGCAGCACCATGATGTGTCGG 1889  
1668 TGGTTTCAATGATGCTCTCTCTGAGGACATCTCTGATTTACCTCGGAACAGTGTGATGC 1727  
1890 TGGTTTCAATGATGCTCTCTCTGAGGACATCTCTGATTTACCTCGGAACAGTGTGATGC 1949

QY 1728 GCATCGAGTGAAGCCGCTTCTTTAGCAGCAGGACCTGATATTCGCTGCTGTGTGGACAC 1787  
Db 1950 GCACTCGAGTGAAGCCGCTTCTTTAGCAGCAGGACCTGATATTCGCTGCTGTGTGGACAC 2009  
QY 1788 AGGTCGCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAGAAAAGTTTAAA 1847  
Db 2010 AGGTCGCTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAGAAAAGTTTAAA 2069  
QY 1848 ATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATTG 1907  
Db 2070 ATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACAGCACACAGATTG 2129  
QY 1908 TTACAGCTGACAGCCCAACCAATGACTGCCACTGGTGCATGCAATGATGTGTCCCCAG 1967  
Db 2130 TTACAGCTGACAGCCCAACCAATGACTGCCACTGGTGCATGCAATGATGTGTCCCCAG 2189  
QY 1968 GAACACAGCTGCTCAGAGGCCAGATCTCCATTTTTTAGGTATGAGATTTGCCCCCAAGGA 2027  
Db 2190 GAACACAGCTGCTCAGAGGCCAGATCTCCATTTTTTAGGTATGAGATTTGCCCCCAAGGA 2249  
QY 2028 TAAACCTATGACTACTGTAAACAAGAGACAGCTGAGGAGCTGTGCCCTGGACCCAGAA 2087  
Db 2250 TAAACCTATGACTACTGTAAACAAGAGACAGCTGAGGAGCTGTGCCCTGGACCCAGAA 2309  
QY 2088 CTGCCAGTGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGCCCGAAAATATCTGTGGCAT 2147  
Db 2310 CTGCCAGTGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGCCCGAAAATATCTGTGGCAT 2369  
QY 2148 TGGCTGGCATTTGTTGGAACTCATGTTTGAATAATTTACTACTGCCAAGGAGATTTATGA 2207  
Db 2370 TGGCTGGCATTTGTTGGAACTCATGTTTGAATAATTTACTACTGCCAAGGAGATTTATGA 2429  
QY 2208 CAATGCTAAAATGTTTCTGTAGGAAACCAATGCCCCCTTTTGGCTTCTCTTCAACCCAGAA 2267  
Db 2430 CAATGCTAAAATGTTTCTGTAGGAAACCAATGCCCCCTTTTGGCTTCTCTTCAACCCAGAA 2489  
QY 2268 GAAGTAGAATTTGTCTTAAAGAGCTGCGAATTAATGCAATCATCTCAGACATGTCCAA 2327  
Db 2490 GAAGTAGAATTTGTCTTAAAGAGCTGCGAATTAATGCAATCATCTCAGACATGTCCAA 2549  
QY 2328 GCTACCTTAACCCATGGGTGCGCTTGGGAAGTCAATGTGCTCTACTGCTGTGGGA 2387  
Db 2550 GCTACCTTAACCCATGGGTGCGCTTGGGAAGTCAATGTGCTCTACTGCTGTGGGA 2609  
QY 2388 AGATATGTCCTCAATTTACAAATAGTTTACTACAGTGAATGCCCTGTGAGCCAGTATGC 2447  
Db 2610 AGATATGTCCTCAATTTACAAATAGTTTACTACAGTGAATGCCCTGTGAGCCAGTATGC 2669  
QY 2448 TGGATTTCTGGAATTTTATCAGAACCCAGTACTCGGGACTGGAAGGTGCAACTGCTCAT 2507  
Db 2670 TGGATTTCTGGAATTTTATCAGAACCCAGTACTCGGGACTGGAAGGTGCAACTGCTCAT 2729  
QY 2508 CAACCCACTCAATGCTAGTGTCTGTGAAGGCTGCAAAACCAACAGTGTCTAAGCAGTGC 2567  
Db 2730 CAACCCACTCAATGCTAGTGTCTGTGAAGGCTGCAAAACCAACAGTGTCTAAGCAGTGC 2789  
QY 2568 GACACCATGTCCTTGAAGCAGCATGTGGAGATTTGCACAGCCGACGCTCTGAGTGCAT 2627  
Db 2790 GACACCATGTCCTTGAAGCAGCATGTGGAGATTTGCACAGCCGACGCTCTGAGTGCAT 2849  
QY 2628 TGGTGGCAGCAACATGAAGCAGTGTGTGGAATCCCAATGCCCTATGTGGCTCTCTTCCCTTT 2687  
Db 2850 GTGGTGGCAGCAACATGAAGCAGTGTGTGGAATCCCAATGCCCTATGTGGCTCTCTTCCCTTT 2909  
QY 2688 TGGCCAGTGTATGGAATGGTATAGATGAGCACTGCCCTCCCTGAAAATTTGTTAGGGTA 2747  
Db 2910 TGGCCAGTGTATGGAATGGTATAGATGAGCACTGCCCTCCCTGAAAATTTGTTAGGGTA 2969  
QY 2748 CTGTACCTGTAGTCAATCTTGGAGCAACAGGCTGTGGCTGGTGTACTGTATCCAGCAA 2807  
Db 2970 CTGTACCTGTAGTCAATCTTGGAGCAACAGGCTGTGGCTGGTGTACTGTATCCAGCAA 3029

QY 2808 TACTGGCAAGGGAATGCAATAGAGGTTCTCTATAAGGACCAAGTGAAGTGCCTTCGCA 2867  
Db 3030 TACTGGCAAGGGAATGCAATAGAGGTTCTCTATAAGGACCAAGTGAAGTGCCTTCGCA 3089  
QY 2868 AGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGA 2927  
Db 3090 AGCCCTTACAGGAAATTTCTATCCACAGCCCTGCTCAATTTCCAGCATGTGTCTAGAGGA 3149  
QY 2928 CAGCAGATACAACTGTGTTCTATCACTGTCCAGCTTGCATATGCAACGGCCACAGTAA 2987  
Db 3150 CAGCAGATACAACTGTGTTCTATCACTGTCCAGCTTGCATATGCAACGGCCACAGTAA 3209  
QY 2988 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGA 3047  
Db 3210 ATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAAGCACTGCGA 3269  
QY 3048 GACCTGCATATCTGGCTTCTACGGTGATCCCAATGGAGGGAATGTGACCATGCAA 3107  
Db 3270 GACCTGCATATCTGGCTTCTACGGTGATCCCAATGGAGGGAATGTGACCATGCAA 3329  
QY 3108 GTGCATGGGACGCTCTGTGACACACACGCGCAAGTGTCTGACACACCAA 3167  
Db 3330 GTGCATGGGACGCTCTGTGACACACACGCGCAAGTGTCTGACACACCAA 3389  
QY 3168 GGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTAGAATAATCGATACCAAGGAAACCC 3227  
Db 3390 GGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTAGAATAATCGATACCAAGGAAACCC 3449  
QY 3228 TCTCAGAGGAACATGTTATTAATCTCTTATGACTATCAGTTCACTTTAGTCTATC 3287  
Db 3450 TCTCAGAGGAACATGTTATTAATCTCTTATGACTATCAGTTCACTTTAGTCTATC 3509  
QY 3288 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGTGTACTCTGACGAAACAAA 3347  
Db 3510 CCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGGTGTACTCTGACGAAACAAA 3569  
QY 3348 CAGGATTTGGACATGTTCAATCAATGCTTCAAGAAATTCACCTCAACATCACCTGGGC 3407  
Db 3570 CAGGATTTGGACATGTTCAATCAATGCTTCAAGAAATTCACCTCAACATCACCTGGGC 3629  
QY 3408 TGCAGTTTCTCAGCTGGACCCAGCTGGAGAGAGTGCCTGTGTTTCAAAACCAA 3467  
Db 3630 TGCAGTTTCTCAGCTGGACCCAGCTGGAGAGAGTGCCTGTGTTTCAAAACCAA 3689  
QY 3468 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCAACCCAAA 3527  
Db 3690 CATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCAACCCAAA 3749  
QY 3528 TATCACTTTCTGTTTATGTCAGTAATTCACCTGGCCCATCAAAATTCAGATTGCCCT 3587  
Db 3750 TATCACTTTCTGTTTATGTCAGTAATTCACCTGGCCCATCAAAATTCAGATTGCCCT 3809  
QY 3588 CTCTCAGCAGCAAAATTTATGACCTGTGACAGTTCTTCTGACATTTCTTCAGTTGTTT 3647  
Db 3810 CTCTCAGCAGCAAAATTTATGACCTGTGACAGTTCTTCTGACATTTCTTCAGTTGTTT 3869  
QY 3648 CCTCTTTCTGCTGCTGCTGCTGTTTGGAGATCAACAAAGTTGTTGGCCCTC 3707  
Db 3870 CCTCTTTCTGCTGCTGCTGCTGTTTGGAGATCAACAAAGTTGTTGGCCCTC 3929  
QY 3708 CAGACGTAGAGAGCAACTTCTTCAGAGATGCAACAGATGGCCGCTCCCTTTGCCCTC 3767  
Db 3930 CAGACGTAGAGAGCAACTTCTTCAGAGATGCAACAGATGGCCGCTCCCTTTGCCCTC 3989  
QY 3768 TGTAATGTGCTTGGAAACAGATGAGGAGCTCTGATCTATTATGGGGGAGATATAA 3827  
Db 3990 TGTAATGTGCTTGGAAACAGATGAGGAGCTCTGATCTATTATGGGGGAGATATAA 4049  
QY 3828 GACTGTTCCCAACCACTTCACTGAGCGCTGTTTGGCAACAAAGCGCTGTCTCTC 3887  
Db 4050 GACTGTTCCCAACCACTTCACTGAGCGCTGTTTGGCAACAAAGCGCTGTCTCTC 4109  
QY 3888 TGTGTTTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCTCTGCGCAGTCAAGTCT 3947

Db 4110 TGTGTTTGTGAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCTGCGCAGTCAAGTCT 4169  
QY 3948 TGTGTTGCGCCACGCGCTGTTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAGGAGAA 4007  
Db 4170 TGTGTTGCGCCACGCGCTGTTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAGGAGAA 4229  
QY 4008 GTCAGGAGCCTGTGAGAACCGGAGACGACGACGCGCTGTCACAGCTGGGACCTGCATCTG 4067  
Db 4230 GTCAGGAGCCTGTGAGAACCGGAGACGACGACGCGCTGTCACAGCTGGGACCTGCATCTG 4289  
QY 4068 A 4068  
Db 4290 A 4290  
RESULT 4  
AAS72659  
ID AAS72659 standard; cDNA; 8561 BP.  
XX AAS72659;  
AC AAS72659;  
XX 13-FEB-2002 (first entry)  
XX DNA encoding novel human diagnostic protein #8463.  
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.  
XX WO200175067-A2.  
XX 11-OCT-2001.  
XX 30-MAR-2001; 2001WO-US008631.  
XX 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX (HYSE-) HYSEQ INC.  
XX Drmanac RT, Liu C, Tang YT;  
XX WPI; 2001-639362/73.  
DR P-PSDB; ABG08472.  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX Claim 1; SEQ ID NO 8463; 103pp; English.  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: the sequence data for this  
CC patent did not appear in the printed specification, but was obtained in



CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

Query Match 97.9%; Score 3981; DB 5; Length 8561;  
Best Local Similarity 99.4%; Pred. No. 0;  
Matches 3996; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 48 GCGCGGACGGAGCGCTCGCGGACGAGCGCGCGCGCACTGTGTCAACGGCGGTG 107  
DB 270 GTCCGGCTCAGCGGACGAGCGGCGCAAGAAATGTGACCGGCTGTGTCAACGGCGGTG 329  
QY 108 CTGCAACCTGGCAACCGGCGAGTGGTCTGCCCGCGCGCTGGTGGGCGAGCAATGCCA 167  
DB 330 CTGCAACCTGGCAACCGGCGAGTGGTCTGCCCGCGCGCTGGTGGGCGAGCAATGCCA 389  
QY 168 GCATCGCGGGCGGCTTCAGACTAATCTGATCTTCTGGGTTGTGACAGATGGACCTGG 227  
DB 390 GCATCGCGGGCGGCTTCAGACTAATCTGATCTTCTGGGTTGTGACAGATGGACCTGG 449  
QY 228 AAAATTATAATACAAACGAAGTGCACTGGCTCAATTGAAGGACAGCCAAATAGAAATAT 287  
DB 450 AAAATTATAATACAAACGAAGTGCACTGGCTCAATTGAAGGACAGCCAAATAGAAATAT 509  
QY 288 GAGACTTCGTTCAATCATTTTGTCTACAGAGTGTAGTTGGGACCAATTTATATGTTATGA 347  
DB 510 GAGACTTCGTTCAATCATTTTGTCTACAGAGTGTAGTTGGGACCAATTTATATGTTATGA 569  
QY 348 TGGGCACTCAATTTATGCAACCGCTAGTTGCTGCATTTAGTGGCTCAATTTCTCGAGAG 407  
DB 570 TGGGCACTCAATTTATGCAACCGCTAGTTGCTGCATTTAGTGGCTCAATTTCTCGAGAG 629  
QY 408 AGATGCAATGAGACTGTCCCTGAGTTGTGGCCATCAGAGTTATGCCCTGCTGATTT 467  
DB 630 AGATGCAATGAGACTGTCCCTGAGTTGTGGCCATCAGAGTTATGCCCTGCTGATTT 689  
QY 468 TTTTGTGATGCTGCTTAAATTTGACTGATTTATATTTACTTACAGTTTGTGATG 527  
DB 690 TTTTGTGATGCTGCTTAAATTTGACTGATTTATATTTACTTACAGTTTGTGATG 749  
QY 528 TCCAAATAACTGCTCAGGCGGAGGAGTGTAAAGATCAGTAATAGCAGGATCTCTGTTGA 587  
DB 750 TCCAAATAACTGCTCAGGCGGAGGAGTGTAAAGATCAGTAATAGCAGGAACTGTTGA 809  
QY 588 ATGTGAATGTTCTGAAAATGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 647  
DB 810 ATGTGAATGTTCTGAAAATGGAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAA 869  
QY 648 CTGTGTTTTCCTCATCGAGGCACTCGCAATTCAGTGTATGTCAGAGGATGCTCCTGCTT 707  
DB 870 CTGTGTTTTCCTCATCGAGGCACTCGCAATTCAGTGTATGTCAGAGGATGCTCCTGCTT 929  
QY 708 CTCAGACTGCGAGGCTCTGGATGTTTCAGTTCCCTGTTACAGCTAACAGTCAITTTGGAC 767  
DB 930 CTCAGACTGCGAGGCTCTGGATGTTTCAGTTCCCTGTTACAGCTAACAGTCAITTTGGAC 989  
QY 768 TCGAGAGGATATTTCTAACTTTAAAGTCCCGAGGATCTCATPAAAGTGTGCTCAATGG 827  
DB 990 TCGAGAGGATATTTCTAACTTTAAAGTCCCGAGGATCTCATPAAAGTGTGCTCAATGG 1049  
QY 828 AAACATTATGTGGTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGTTCT 887  
DB 1050 AAACATTATGTGGTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGTTCT 1109  
QY 888 AGCGTATGACTTCTTCTAGGAGTGGCTTCACCTAAACCGTTCGTGAACAATGTGCT 947  
DB 1110 AGCGTATGACTTCTTCTAGGAGTGGCTTCACCTAAACCGTTCGTGAACAATGTGCT 1169  
QY 948 TGTATGATATGTTCAATTTCTTTGGCAATTTATCAAGGATAAAATTTACATGTATGGAGAAA 1007  
DB 1170 TGTATGATATGTTCAATTTCTTTGGCAATTTATCAAGGATAAAATTTACATGTATGGAGAAA 1229

QY 1008 AATTGATTCAACTGGGAATGTGACCAATGATGTGAGAGTGTTCACATTCATATGAGTC 1067  
DB 1230 AATTGATCCAACTGGGAATGTGACCAATGATGTGAGAGTGTTCACATTCATATGAGTC 1289  
QY 1068 ATGGGTGCTGTGTGACCCCTTAAGGCAAGAGGAGTATGAGTGTGGGCACTCTGCACA 1127  
DB 1290 ATGGGTGCTGTGTGACCCCTTAAGGCAAGAGGAGTATGAGTGTGGGCACTCTGCACA 1349  
QY 1128 CATTTGTACACTGGAAGAAATGCGCGAGTGGTCACTGCTGGTCACTTTGGTCACTGCCCTCT 1187  
DB 1350 CATTTGTACACTGGAAGAAATGCGCGAGTGGTCACTGCTGGTCACTTTGGTCACTGCCCTCT 1409  
QY 1188 CTATGGATATATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1247  
DB 1410 CTATGGATATATAGCAATGTGAGGAATATGATTTGGATAAGAACACATGGAGTATATT 1469  
QY 1248 ACACACCCAGGCTGCCCTTTGTGAAGGGGGTTACGGCCATAGCAGTGTTCACGCCATAG 1307  
DB 1470 ACACACCCAGGCTGCCCTTTGTGAAGGGGGTTACGGCCATAGCAGTGTTCACGCCATAG 1529  
QY 1308 GACAGGGCCCTATACGTTCTATGCTGGCTACAAGGCTTTTCACTGCCAATAGTACCGCT 1367  
DB 1530 GACAGGGCCCTATACGTTCTATGCTGGCTACAAGGCTTTTCACTGCCAATAGTACCGCT 1589  
QY 1368 TGCAGATGATCTCTACCGATATGATGTGATACCCAGATGTGGACCAATCTTAAAGGACAG 1427  
DB 1590 TGCAGATGATCTCTACCGATATGATGTGATACCCAGATGTGGACCAATCTTAAAGGACAG 1649  
QY 1428 CCGATTTTTCGGTTACTTGTGACACAGCTGTGATAGTGGAGAACCAATGCTGTGTTGG 1487  
DB 1650 CCGATTTTTCGGTTACTTGTGACACAGCTGTGATAGTGGAGAACCAATGCTGTGTTGG 1709  
QY 1488 GGGAAACACACAAATGACATCTATGAGCCATGCGCGCCAAATGCTTCTTTCAGATTT 1547  
DB 1710 GGGAAACACACAAATGACATCTATGAGCCATGCGCGCCAAATGCTTCTTTCAGATTT 1769  
QY 1548 CATGCCCTATGACATTTGCTGTGACCGCTGTGAGTGTCTTCCAGACCTGATCTCCACCA 1607  
DB 1770 CATGCCCTATGACATTTGCTGTGACCGCTGTGAGTGTCTTCCAGACCTGATCTCCACCA 1829  
QY 1608 TGATGTCAAAGATTTGGCCATTCAGCAGTCTTACAAACAGCAGCAGTATGTTGTCGG 1667  
DB 1830 TGATGTCAAAGATTTGGCCATTCAGCAGTCTTACAAACAGCAGCAGTATGTTGTCGG 1889  
QY 1668 TGGTTTCAATAGTCTCTCTCCTCAGCAGCATCTCTGGTATTCACCTCGGAAACAGTGTGATC 1727  
DB 1890 TGGTTTCAATAGTCTCTCTCCTCAGCAGCATCTCTGGTATTCACCTCGGAAACAGTGTGATC 1949  
QY 1728 GCATCGGAGTGAAGCGGCTTGTGACAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 1787  
DB 1950 GCATCGGAGTGAAGCGGCTTGTGACAGCAGGACCTGGTATTCGGTGTGTGTGGAACAC 2009  
QY 1788 AGGTCGTCTCAGTGTATCTGTCGGCGCTGGCAACTGATGAAACAGAGAAAGTTAAA 1847  
DB 2010 AGGTCGTCTCAGTGTATCTGTCGGCGCTGGCAACTGATGAAACAGAGAAAGTTAAA 2069  
QY 1848 ATCAGATGTTTTCAAAAGAACTCTTTGACCATGACAGATGTGACAGCAGACAGATTG 2129  
DB 2070 ATCAGATGTTTTCAAAAGAACTCTTTGACCATGACAGATGTGACAGCAGACAGATTG 2129  
QY 1908 TTACAGCTGCACAGCCCAACCAATGACTGCCATGCTGGTGAATGCAATGCTGTGTCCCGAG 1967  
DB 2130 TTACAGCTGTACAGCCCAACCAATGACTGCCATGCTGGTGAATGCAATGCTGTGTCCCGAG 2189  
QY 1968 GAACACAGCTGTCTCAGAAAGCCAGATCTCCATTTTTAGGTATGAGAAATTCGCCCAAGGA 2027  
DB 2190 GAACACAGCTGTCTCAGAAAGCCAGATCTCCATTTTTAGGTATGAGAAATTCGCCCAAGGA 2249  
QY 2028 TAAACCTATGACTACTGTAAACAGAGACAGCTGAGGAGCTGTGCCCTGGGACAGAA 2087  
DB 2250 TAAACCTATGACTACTGTAAACAGAGACAGCTGAGGAGCTGTGCCCTGGGACAGAA 2309  
QY 2088 CTGCGAGTGGAGCCCGCGGAATCAGGAGTGCATTTGCCCTGCTGCCGAAAAATATCTGTGGCAT 2147



Db	2310	CTGCCAGTGGGAGCCCGGAATCAGGAGTCAATGGCCCTGCCCGGAAAAATATCTGTGGCAT	2369
Qy	2148	TGGCTGCATTTGGTTGGAAACTCATGTTTGAATAATTAATCTACTGCGCAAGAGAAATTATGA	2207
Db	2370	TGGCTGCATTTGGTTGGAAACTCATGTTTGAATAATTAATCTACTGCGCAAGAGAAATTATGA	2429
Qy	2208	CAATGCTAAATTTGTTCTGTAGGAACCAATGCGCTTTTGGCTTCTCTTACAAACCCAGAA	2267
Db	2430	CAATGCTAAATTTGTTCTGTAGGAACCAATGCGCTTTTGGCTTCTCTTACAAACCCAGAA	2489
Qy	2268	GAGGTTAGAAATTTGTTCTTAAAGCAGTGGCAATAAATGCAATGATCTACAGAGCAATGCCAA	2327
Db	2490	GAGGTTAGAAATTTGTTCTTAAAGCAGTGGCAATAAATGCAATGATCTACAGAGCAATGCCAA	2549
Qy	2328	GCTCACCTTAACCCATGGGTGGGCTTGGGAAGATCAATGCTCTCTACTGTTGGTGGGA	2387
Db	2550	GCTCACCTTAACCCATGGGTGGGCTTGGGAAGATCAATGCTCTCTACTGTTGGTGGGA	2609
Qy	2388	AGATATGTCCTTACAAATAGTTTACTACAGTGGATGCCCTCTGAGCCCAAGTATGC	2447
Db	2610	AGATATGTCCTTACAAATAGTTTACTACAGTGGATGCCCTCTGAGCCCAAGTATGC	2669
Qy	2448	TGGATTTCTGTGGAAATTTATCAGAACCCAGTACTCGGGNCTGAAGCTGCAACCTGCAT	2507
Db	2670	TGGATTTCTGTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCAT	2729
Qy	2508	CAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCAAGTCTGAAGCAGTGGCG	2567
Db	2730	CAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCAAGTCTGAAGCAGTGGCG	2789
Qy	2568	GACACCAATGTCCTTGAGGACAGATGTGAGATTGCAACAGGCGCAGCTCTGAGTGCAT	2627
Db	2790	GACACCAATGTCCTTGAGGACAGATGTGAGATTGCAACAGGCGCAGCTCTGAGTGCAT	2849
Qy	2628	GTGTTGACAGCAATGAAGCAGTGTGGACTCCAATGCTATGTGGCTCTCTTCCCTTT	2687
Db	2850	GTGTTGACAGCAATGAAGCAGTGTGGACTCCAATGCTATGTGGCTCTCTTCCCTTT	2909
Qy	2688	TGCGCAATGTATGGAATGATACGATAGACACCTGCCCTTGAATAATGTTTACGGCTA	2747
Db	2910	TGCGCAATGTATGGAATGATACGATAGACACCTGCCCTTGAATAATGTTTACGGCTA	2969
Qy	2748	CTGTACCTGTAGTATGTTGAGCAACAGAGCTGTGGTGTACTGTACTGTCTAGAGGA	2807
Db	2970	CTGTACCTGTAGTATGTTGAGCAACAGAGCTGTGGTGTACTGTACTGTCTAGAGGA	3029
Qy	2808	TACTGGCAAGGAAATGCAATAGAGGTTTCTTATAAGGACCAAGTGAAGATGCTTTCGCA	2867
Db	3030	TACTGGCAAGGAAATGCAATAGAGGTTTCTTATAAGGACCAAGTGAAGATGCTTTCGCA	3089
Qy	2868	AGCCCTTACAGAAATTTCTATCCAGCCCTGCTCAATTCAGCATGTGTCTAGAGGA	2927
Db	3090	AGCCCTTACAGAAATTTCTATCCAGCCCTGCTCAATTCAGCATGTGTCTAGAGGA	3149
Qy	2928	CAGCAGATACAATGTGCTTTCTATCTGTCAGCTTGCATGCAACCGGCAACAGTAA	2987
Db	3150	CAGCAGATACAATGTGCTTTCTATCTGTCAGCTTGCATGCAACCGGCAACAGTAA	3209
Qy	2988	ATGATCAATCAGAGCATCTGTAGAGATGTGAGAACCTTGACCAAGGCAAGCATGGGA	3047
Db	3210	ATGATCAATCAGAGCATCTGTAGAGATGTGAGAACCTTGACCAAGGCAAGCATGGGA	3269
Qy	3048	GACTGTATCTGTGGCTTCTACCGTGTATCCCAATGAGGGAATGTCAAGCATGCA	3107
Db	3270	GACTGTATCTGTGGCTTCTACCGTGTATCCCAATGAGGGAATGTCAAGCATGCA	3329
Qy	3108	GTGCAATGGGCAACCGGCTCTGTGCAACCAACAGCGGCAAGTGTCTGCAACCA	3167
Db	3330	GTGCAATGGGCAACCGGCTCTGTGCAACCAACAGCGGCAAGTGTCTGCAACCA	3389
Qy	3168	GGCGCTCAAGGGGAGGAGTGCAGTATGTGAGGTAGAAAATCGATATCAAGGAAACCC	3227

Db	3390	GGCGCTCAAGGGGAGGAGTGCAGTATGTGAGGTAGAAAATCGATATCAAGGAAACCC	3449
Qy	3228	TCTCAGAGGACATGTTATTAATCTCTTATTGACTATCAGTTCACTTTAGTCTATC	3287
Db	3450	TCTCAGAGGACATGTTATTAATCTCTTATTGACTATCAGTTCACTTTAGTCTATC	3509
Qy	3288	CCAGGACATGATCGCTATTACACAGCTATCAAATTTTGTGGCTACTCTGACGAACAAA	3347
Db	3510	CCAGGACATGATCGCTATTACACAGCTATCAAATTTTGTGGCTACTCTGACGAACAAA	3569
Qy	3348	CAGGATTTGACATGTTTCATCAATGCTCCAAAGATTTCAACCTCAACATCACTGGGC	3407
Db	3570	CAGGATTTGACATGTTTCATCAATGCTCCAAAGATTTCAACCTCAACATCACTGGGC	3629
Qy	3408	TGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCA	3467
Db	3630	TGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCA	3689
Qy	3468	CATTAAAGAGTACAAAGATGTTTCTCTAATGAGAAGTTTGAATTTTCGCAACCAACAAA	3527
Db	3690	CATTAAAGAGTACAAAGATGTTTCTCTAATGAGAAGTTTGAATTTTCGCAACCAACAAA	3749
Qy	3528	TATCACTTTCTTCTGTTTATGTCAGTAAATTCACCTGCCCATCAAAATTCAGATTGCTT	3587
Db	3750	TATCACTTTCTTCTGTTTATGTCAGTAAATTCACCTGCCCATCAAAATTCAGATTGCTT	3809
Qy	3588	CTCTCAGCACAGCAATTTATGGAACCTGGTACAGTTCTTCTGAGCTTTCTTCAAGTTGTTT	3647
Db	3810	CTCTCAGCACAGCAATTTATGGAACCTGGTACAGTTCTTCTGAGCTTTCTTCAAGTTGTTT	3869
Qy	3648	CCTCTTTTCTCTGCTGGTGGCTGCTGTTTGGAGATCAAAAGTTGTTGGGCTC	3707
Db	3870	CCTCTTTTCTCTGCTGGTGGCTGCTGTTTGGAGATCAAAAGTTGTTGGGCTC	3929
Qy	3708	CAGACGTAGAGAGCAACTTCTTCCAGAGATGCAACAGATGCGCCGCTTCTTGGCTC	3767
Db	3930	CAGACGTAGAGAGCAACTTCTTCCAGAGATGCAACAGATGCGCCGCTTCTTGGCTC	3989
Qy	3768	TGTAATGTGCTTGGAAACAGATGAGGAGCTTCTGATCTTATTGGGGGAGTATAAA	3827
Db	3990	TGTAATGTGCTTGGAAACAGATGAGGAGCTTCTGATCTTATTGGGGGAGTATAAA	4049
Qy	3828	GACTGTTTCCCAAAACCCATTGCACTGGAGCCGTGTTTGGCAACAAAGCCGCTCTCTC	3887
Db	4050	GACTGTTTCCCAAAACCCATTGCACTGGAGCCGTGTTTGGCAACAAAGCCGCTCTCTC	4109
Qy	3888	TGTTGTTGAGGCTCCCTCGAGCCCTGGGTGGCATCCCTCTCTGGGAGTCAAGTCT	3947
Db	4110	TGTTGTTGAGGCTCCCTCGAGCCCTGGGTGGCATCCCTCTCTGGGAGTCAAGTCT	4169
Qy	3948	TGCTGTGGCCAGCCCTCGGTGGACATTTCTCAGCAGATCCGATAGTGTACAGAGAA	4007
Db	4170	TGCTGTGGCCAGCCCTCGGTGGACATTTCTCAGCAGATCCGATAGTGTACAGAGAA	4229
Qy	4008	GTCCAGGAGCGGTGAGAAACCGGAGAGAGAGCCCTTGCACAGCCTGGGACCTGCACTG	4067
Db	4230	GTCCAGGAGCGGTGAGAAACCGGAGAGAGAGCCCTTGCACAGCCTGGGACCTGCACTG	4289
Qy	4068	A 4068	
Db	4290	A 4290	

RESULT 5  
AAZ52274  
ID AAZ52274 standard; cDNA; 3597 BP.  
XX  
AC AAZ52274;  
XX  
DT 18-JUL-2000 (first entry)  
XX  
DE Human soluble attractin-1 cDNA.  
XX



Db 1381 TACCGATATGATGTGGATACCCAGATGTGGACCAATCTTAAGGACAGCCGATTTTCCGT 1440  
Qy 1441 TACTTGCACACAGCTGTGATGTAGTGGAAACCATGCTGCTGTTTGGGGGAACACACAC 1500  
Db 1441 TACTTGCACACAGCTGTGATGTAGTGGAAACCATGCTGCTGTTTGGGGGAACACACAC 1500  
Qy 1501 AATGACACATCTATGACCATGGCCCAATGCTTCTCTTTCAGATTTTCATGGCCTATGAC 1560  
Db 1501 AATGACACATCTATGACCATGGCCCAATGCTTCTCTTTCAGATTTTCATGGCCTATGAC 1560  
Qy 1561 ATTGCTGTGACCGCTGTGATGTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGA 1620  
Db 1561 ATTGCTGTGACCGCTGTGATGTGCTTCCAGACCTGAT-TCACACCATGATGTCAACAGA 1619  
Qy 1621 TTTGGCCATT-CAGCAGTCTTACCAACAGCACCATGATGTGTTGGTGGTTTCAATAG 1679  
Db 1620 TTTGGCCATTCCAGCAGCTTTACACACAGACCAATGATGTGTTGGTGGTTTCAATAG 1679  
Qy 1680 TCTCCTCCTCAGCAGCATCCTGTTATTCACCTCGAAACAGTGTGATGCGCATCGAGTGA 1739  
Db 1680 TCTCCTCCTCAGCAGCATCCTGTTATTCACCTCGAAACAGTGTGATGCGCATCGAGTGA 1739  
Qy 1740 AGCCGCTTGTAGCAGCAGCAGCTGTTATTCGCTGTGCTGGAACACAGGCTGCTCA 1799  
Db 1740 AGCCGCTTGTAGCAGCAGCAGCTGTTATTCGCTGTGCTGGAACACAGGCTGCTCA 1799  
Qy 1800 GTGTATCTCGTGGCGCTGCAACTGATGAACAAAGAAAGTTAAAAATCAGAAATGTTT 1859  
Db 1800 GTGTATCTCGTGGCGCTGCAACTGATGAACAAAGAAAGTTAAAAATCAGAAATGTTT 1859  
Qy 1860 TTCCAAAGAACTTTGACACAGCAGCAGTGTGACCAAGCAGACACAGATGTTTACAGCTGCAC 1919  
Db 1860 TTCCAAAGAACTTTGACACAGCAGCAGTGTGACCAAGCAGACACAGATGTTTACAGCTGCAC 1919  
Qy 1920 AGCCAAACCAATGACTGCACTGCTGCAATGACCAATGCTGCTCCAGGAACACACAGCTG 1979  
Db 1920 AGCCAAACCAATGACTGCACTGCTGCAATGACCAATGCTGCTCCAGGAACACACAGCTG 1979  
Qy 1980 CTGAGAGGCGAGATCTCCATTTTGTAGGATGAGAAATGCGCCCAAGATTAACCTATGTA 2039  
Db 1980 CTGAGAGGCGAGATCTCCATTTTGTAGGATGAGAAATGCGCCCAAGATTAACCTATGTA 2039  
Qy 2040 CTACTGTAAACAAGAACACAGCTGACAGAGCTGTGCTGACAGACCACTGCCAGTGGGA 2099  
Db 2040 CTACTGTAAACAAGAACACAGCTGACAGAGCTGTGCTGACAGACCACTGCCAGTGGGA 2099  
Qy 2100 GCCCCGAATCAGAGGTGCAATGCTGCTGCCGAAATATCTGTGGCAATGGCTGGCAATTT 2159  
Db 2100 GCCCCGAATCAGAGGTGCAATGCTGCTGCCGAAATATCTGTGGCAATGGCTGGCAATTT 2159  
Qy 2160 GGTGTGAAATCTCATGTTTGAATTTACTGCTGCAAGGAAATATGACAAATGCTAAAT 2219  
Db 2160 GGTGTGAAATCTCATGTTTGAATTTACTGCTGCAAGGAAATATGACAAATGCTAAAT 2219  
Qy 2220 GTTCTGTAGGAACCAATGCCCTTTTGGCTTCTCTTACACCCAGAAGAGGTAGAAT 2279  
Db 2220 GTTCTGTAGGAACCAATGCCCTTTTGGCTTCTCTTACACCCAGAAGAGGTAGAAT 2279  
Qy 2280 TGTCTTAAAGCAGTGGAAATATGATCATCTCAGAGCATGTCGAAGCTCACCTTAC 2339  
Db 2280 TGTCTTAAAGCAGTGGAAATATGATCATCTCAGAGCATGTCGAAGCTCACCTTAC 2339  
Qy 2340 CCCATGGCTGGCCTTCCGAAGATCAATGTGCTTCTACTGCTGGGAAGATATGTCCTCC 2399  
Db 2340 CCCATGGCTGGCCTTCCGAAGATCAATGTGCTTCTACTGCTGGGAAGATATGTCCTCC 2399  
Qy 2400 ATTACAAATAGTTTACTACAGTGGATGCGCTGTGAGCCAGTGTGCTGGATTCTGTGG 2459  
Db 2400 ATTACAAATAGTTTACTACAGTGGATGCGCTGTGAGCCAGTGTGCTGGATTCTGTGG 2459  
Qy 2460 AATTTTATCAGAACCCAGTACTCGGGACTGAAGGCTGAACCTGCAATCAACCCACTCAA 2519  
Db 2460 AATTTTATCAGAACCCAGTACTCGGGACTGAAGGCTGAACCTGCAATCAACCCACTCAA 2519

Qy 2520 TGGTAGTCTGTGTAAAGGCTGCAAAACCAAGTGTCTAAGCAGTGCAGACCAACCATGTGC 2579  
Db 2520 TGGTAGTCTGTGTAAAGGCTGCAAAACCAAGTGTCTAAGCAGTGCAGACCAACCATGTGC 2579  
Qy 2580 CTTGAGGACAGCATGTGAGATTTGACACAGCGGAGCTCTGAGTGCATGTGTGTCAGCAA 2639  
Db 2580 CTTGAGGACAGCATGTGAGATTTGACACAGCGGAGCTCTGAGTGCATGTGTGTCAGCAA 2639  
Qy 2640 CATGAAGCAGTGTGTGAGTCCATGCTATGTGSCCTCCTTCCCTTTTGGCCAGTGTAT 2699  
Db 2640 CATGAAGCAGTGTGTGAGTCCATGCTATGTGSCCTCCTTCCCTTTTGGCCAGTGTAT 2699  
Qy 2700 GGAATGATATACGATGAGACCTGCCCTCCCTGAAAAATTTGTTTCAGGCTACTGTACTGTAG 2759  
Db 2700 GGAATGATATACGATGAGACCTGCCCTCCCTGAAAAATTTGTTTCAGGCTACTGTACTGTAG 2759  
Qy 2760 TCATTGCTTGGAGCAACAGGCTGTGCTGCTGATCTGATCCAGCAATACTTGGCAAGG 2819  
Db 2760 TCATTGCTTGGAGCAACAGGCTGTGCTGCTGATCTGATCCAGCAATACTTGGCAAGG 2819  
Qy 2820 GAAATGCATAGAGGCTTCTTATAAAGGACCAAGTCAAGATGCTTCGCAAGCCCTACAGG 2879  
Db 2820 GAAATGCATAGAGGCTTCTTATAAAGGACCAAGTCAAGATGCTTCGCAAGCCCTACAGG 2879  
Qy 2880 AAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAA 2939  
Db 2880 AAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAA 2939  
Qy 2940 CTGCTTCTTCACTTCACTGCTCCAGCTTGCATGCAAGCCGACACAGTAAATGCAATCA 2999  
Db 2940 CTGCTTCTTCACTTCACTGCTCCAGCTTGCATGCAAGCCGACACAGTAAATGCAATCA 2999  
Qy 3000 GAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAAGCAGTGCAGACCTTGCATATC 3059  
Db 3000 GAGCATCTGTGAGAGTGTGAGAACCTGACACAGGCAAGCAGTGCAGACCTTGCATATC 3059  
Qy 3060 TGGCTTCTACGCTGATCCCAATGAGGGAATGTGCGCCATGCAAGTGCATGCGCAATGGCA 3119  
Db 3060 TGGCTTCTACGCTGATCCCAATGAGGGAATGTGCGCCATGCAAGTGCATGCGCAATGGCA 3119  
Qy 3120 CGCGTCTCTGTGCAACCAACACGCGGCAAGTGTCTTGCACCAACCAAGGCGTCAAGGG 3179  
Db 3120 CGCGTCTCTGTGCAACCAACACGCGGCAAGTGTCTTGCACCAACCAAGGCGTCAAGGG 3179  
Qy 3180 GGAAGAGTCCAGTATGTGAGGTAGAAATCGATCCAGGAACCTCTCAGAGGAAC 3239  
Db 3180 GGAAGAGTCCAGTATGTGAGGTAGAAATCGATCCAGGAACCTCTCAGAGGAAC 3239  
Qy 3240 ATGTTATTTATCTCTTCTTATTTAGTATCAGTTTCACTTTAGTCTATCCAGGAAGATGA 3299  
Db 3240 ATGTTATTTATCTCTTCTTATTTAGTATCAGTTTCACTTTAGTCTATCCAGGAAGATGA 3299  
Qy 3300 TCGCTATTTACACAGTATCAATTTTGTGGCTACTCTCTGACGAAACCAACAGGATTTGGA 3359  
Db 3300 TCGCTATTTACACAGTATCAATTTTGTGGCTACTCTCTGACGAAACCAACAGGATTTGGA 3359  
Qy 3360 CATGTTTCAATCAATGCTCCAAAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTCTC 3419  
Db 3360 CATGTTTCAATCAATGCTCCAAAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTCTC 3419  
Qy 3420 AGCTGGAACCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAACATTAAAGGATGA 3479  
Db 3420 AGCTGGAACCCAGGCTGGAGAGATGCTGTGTTTCAAAAACCAACATTAAAGGATGA 3479  
Qy 3480 CAAGATAGTTTCTTCAATGAGAGTTTGAATTTTCGCAACACCCCAATATCACTTTCTT 3539  
Db 3480 CAAGATAGTTTCTTCAATGAGAGTTTGAATTTTCGCAACACCCCAATATCACTTTCTT 3539  
Qy 3540 TGTATTATGTCAGTAATTTTCACTTGGCCCATCAAAATTCAGAT 3581  
Db 3540 TGTATTATGTCAGTAATTTTCACTTGGCCCATCAAAATTCAGAT 3581

## RESULT 6

AA572657  
ID AAS72657 standard; cDNA; 3597 BP.

XX AC AAS72657;

XX DT 13-FEB-2002 (first entry)

XX DE DNA encoding novel human diagnostic protein #8461.

XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX OS Homo sapiens.

XX PN W0200175067-A2.

XX PD 11-OCT-2001.

XX PF 30-MAR-2001; 2001WO-US008631.

XX PR 31-MAR-2000; 2000US-00540217.

XX PR 23-AUG-2000; 2000US-00649167.

XX PA (HYSE-) HYSEQ INC.

XX PI Drmanac RT, Liu C, Tang YT;

XX DR WPI; 2001-639362/73.

XX DR P-PSDB; ABG08470.

XX PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.

XX PS Claim 1; SEQ ID NO 8461; 103pp; English.

XX CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: the sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 3597 BP; 951 A; 822 G; 926 T; 0 U; 0 Other;

Query Match 87.4%; Score 3556.4; DB 5; Length 3597;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3579; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

QY 1 ATGTGTGCGCCGAGCGGGCAACTGAGGCAAGGCTGAGGAGGAGCGCGCGACGGCA 60

DB 1 ATGTGTGCGCCGAGCGGGCAACTGAGGCAAGGCTGAGGAGGAGCGCGCGACGGCA 60

QY 61 GCGTCTCGCGGCGAGGAGCGGGCCGCCACTGTGTCAACGCGGCTCGCTGCACACCTTGGC 120  
|||||

DB 61 GCGTCTCGCGGCGAGGAGCGGGCCGCCACTGTGTCAACGCGGCTCGCTGCACACCTTGGC 120  
QY 121 ACCGCGCAGTGCCTCTGCGCGCGCGCTGGGTGGCGAGCAATGCGACGACTCGGGGGGC 180  
DB 121 ACCGCGCAGTGCCTCTGCGCGCGCGCTGGGTGGCGAGCAATGCGACGACTCGGGGGGC 180  
QY 181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATATATAATAC 240  
DB 181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATATATAATAC 240  
QY 241 AAAACGAGTGCACGCTGGCTCATTTGAAGGACAGCAATAGATAATAGACACTTCGTTTC 300  
DB 241 AAAACGAGTGCACGCTGGCTCATTTGAAGGACAGCAATAGATAATAGACACTTCGTTTC 300  
QY 301 AATCAATTTGCTACAGAGTGTAGTTGGACCAATATATATATATATATATATATATATAT 360  
DB 301 AATCAATTTGCTACAGAGTGTAGTTGGACCAATATATATATATATATATATATATATAT 360  
QY 361 TATGACCGCTAGTTGCTGCANTTTAGTGGCTCATTTGTTCTTCTGAGAGATGGCAATGAG 420  
DB 361 TATGACCGCTAGTTGCTGCANTTTAGTGGCTCATTTGTTCTTCTGAGAGATGGCAATGAG 420  
QY 421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCCTTGTGCTCATTTTTTTAGTATGCT 480  
DB 421 ACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCCCTTGTGCTCATTTTTTTAGTATGCT 480  
QY 481 GCTTATATTTGACTGGATTTAATATTACTTACAGTTTGTGATATGTGTCCAAATAAATGCG 540  
DB 481 GCTTATATTTGACTGGATTTAATATTACTTACAGTTTGTGATATGTGTCCAAATAAATGCG 540  
QY 541 TCAGCCCGAGGAGTGTAGATCAGTAATAGCAGCGATCTGTTGAATGTGAATGTCT 600  
DB 541 TCAGCCCGAGGAGTGTAGATCAGTAATAGCAGCGATCTGTTGAATGTGAATGTCT 600  
QY 601 GAAAACTGAAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCT 660  
DB 601 GAAAACTGAAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCT 660  
QY 661 CATCGAGGCATCTGCAATTCAGTGTAGTGTAGAGATGCTCTCTCTCTCAGACTGGCAG 720  
DB 661 CATCGAGGCATCTGCAATTCAGTGTAGTGTAGAGATGCTCTCTCTCTCAGACTGGCAG 720  
QY 721 GGTCTGTGATGTTTCACTTCTGTACAGTAAACAGTCAATTTTGGACTCGAGAGGAATAT 780  
DB 721 GGTCTGTGATGTTTCACTTCTGTACAGTAAACAGTCAATTTTGGACTCGAGAGGAATAT 780  
QY 781 TCTAACTTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTTATGTGG 840  
DB 781 TCTAACTTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTTATGTGG 840  
QY 841 GTTGTGGAGGATATATGTTTCAACACTCAGATTATACATGTTCTAGCGTATGACCTT 900  
DB 841 GTTGTGGAGGATATATGTTTCAACACTCAGATTATACATGTTCTAGCGTATGACCTT 900  
QY 901 GCTTCTAGGGAGTGGCTTCCCACTAAACCGTTCTGTGAAACAATGTGGTTGTAGATATGCT 960  
DB 901 GCTTCTAGGGAGTGGCTTCCCACTAAACCGTTCTGTGAAACAATGTGGTTGTAGATATGCT 960  
QY 961 CATTTCTTTGGCATTTATACAGGATAAAATTTTACATGTATGGAGGAAAAATTTGATTTCACT 1020  
DB 961 CATTTCTTTGGCATTTATACAGGATAAAATTTTACATGTATGGAGGAAAAATTTGATTTCACT 1020  
QY 1021 GGGAAATGTGACCAATGAGTTGAGAGTGTGTTTTCACATTTCAATATGAGTCAATGGGTGTGTTG 1080  
DB 1021 GGGAAATGTGACCAATGAGTTGAGAGTGTGTTTTCACATTTCAATATGAGTCAATGGGTGTGTTG 1080  
QY 1081 ACCCTTAAGGCAAAAGGAGGAGTATGCACTGTGGGCACTCTGACACATTTGTACACTG 1140  
DB 1081 ACCCTTAAGGCAAAAGGAGGAGTATGCACTGTGGGCACTCTGACACATTTGTACACTG 1140  
QY 1141 AAGAAATGCGCGAGTGGTCACTGTGCTCATTTTGTGCTCACTGCCCTCTCTATGGATATATA 1200  
DB 1141 AAGAAATGCGCGAGTGGTCACTGTGCTCATTTTGTGCTCACTGCCCTCTCTATGGATATATA 1200

QY	1201	AGCAATGTGACGGAATATGATTTGGATAAGAACACATGGAGTATATTTACACACCCAGGGT	1260	2280	TGTCCTTAAGCAGCTGCGAATAATGCAATGATCTCTCAGAGCATGTCCAAAGCTCACCTTAAC	2339
Db	1201	AGCAATGTGACGGAATATGATTTGGATAAGAACACATGGAGTATATTTACACACCCAGGGT	1260	2280	TGTCCTTAAGCAGCTGCGAATAATGCAATGATCTCTCAGAGCATGTCCAAAGCTCACCTTAAC	2339
QY	1261	GCCCTTGTGCAAGGGGTACGGCCATAGCAGTGTGTTACACCATAGGACACGGGCCCTA	1320	2340	CCCATGGGTGCGCCCTTTCGGAAGATCAATGTGCTCTACTGGTGGGAAAGATATGTCCCC	2399
Db	1261	GCCCTTGTGCAAGGGGTACGGCCATAGCAGTGTGTTACACCATAGGACACGGGCCCTA	1320	2340	CCCATGGGTGCGCCCTTTCGGAAGATCAATGTGCTCTACTGGTGGGAAAGATATGTCCCC	2399
QY	1321	TACGTTCAATGCTGCTACAGGCTTTTCAGTGCCAAATAGTACCGGCTTTCAGATGATCTC	1380	2400	ATTTCACAAATAGTTTACTACAGTGGATGCTCTGAGCCCACTCAACCCACTCAA	2459
Db	1321	TACGTTCAATGCTGCTACAGGCTTTTCAGTGCCAAATAGTACCGGCTTTCAGATGATCTC	1380	2400	ATTTCACAAATAGTTTACTACAGTGGATGCTCTGAGCCCACTCAACCCACTCAA	2459
QY	1381	TACCGATATGATGAGGATACCCAGATGTGACCAATCTTTAAGGACAGCCGATTTTCGGT	1440	2460	AAATTTATCAGAACCCAGTACTCGGGGACTGAAAGCTGCAACCTGCAATCAACCCACTCAA	2519
Db	1381	TACCGATATGATGAGGATACCCAGATGTGACCAATCTTTAAGGACAGCCGATTTTCGGT	1440	2460	AAATTTATCAGAACCCAGTACTCGGGGACTGAAAGCTGCAACCTGCAATCAACCCACTCAA	2519
QY	1441	TACTTGACACAGCTGTGATAGTGAAGAACCATGCTGCTGTTTGGGGGAAACACACAC	1500	2520	TGGTAGTGTCTGTGAAAGGCTGCAAAACCAAGTGTCTAAGCAGTSCCGGACACCATGTGC	2579
Db	1441	TACTTGACACAGCTGTGATAGTGAAGAACCATGCTGCTGTTTGGGGGAAACACACAC	1500	2520	TGGTAGTGTCTGTGAAAGGCTGCAAAACCAAGTGTCTAAGCAGTSCCGGACACCATGTGC	2579
QY	1501	AATGACATCTATGAGCCATGCGGCCAAATGCTTCTCTCAGATTTTCAGGCTATGAC	1560	2580	CTTGAGGACAGCATGTGGAGATTGCAACGCGGAGCTCTGAGTGCATGTGGTGCAGCAA	2639
Db	1501	AATGACATCTATGAGCCATGCGGCCAAATGCTTCTCTCAGATTTTCAGGCTATGAC	1560	2580	CTTGAGGACAGCATGTGGAGATTGCAACGCGGAGCTCTGAGTGCATGTGGTGCAGCAA	2639
QY	1561	ATTGCCCTGTGACCGCTGCTGATGAGTGAAGAACCATGCTGCTGTTTGGGGGAAACACACAC	1620	2640	CATGAAGCAGTGTGGAGTCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTAT	2699
Db	1561	ATTGCCCTGTGACCGCTGCTGATGAGTGAAGAACCATGCTGCTGTTTGGGGGAAACACACAC	1620	2640	CATGAAGCAGTGTGGAGTCCCAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTAT	2699
QY	1621	TTTGGCCATT-CAGCAGCTTTACACACACCATGATGTGCTTCCGTTGTTTCAATAG	1679	2700	GGAATGGTATACGATGAGACCTGCGCCCTGCTGCTTCCCTTTTGGCCAGTGTAT	2759
Db	1621	TTTGGCCATT-CAGCAGCTTTACACACACCATGATGTGCTTCCGTTGTTTCAATAG	1679	2700	GGAATGGTATACGATGAGACCTGCGCCCTGCTGCTTCCCTTTTGGCCAGTGTAT	2759
QY	1680	TCTCTCTCAGGACATCTGCTGATTTACCTCGGACAGTGTGATGCGCATCGGAGTGA	1739	2760	TCATTTGCTTGAGCAACACAGGCTGTGGCTGTGTACTGATCCAGCAATATCTGCAAGG	2819
Db	1680	TCTCTCTCAGGACATCTGCTGATTTACCTCGGACAGTGTGATGCGCATCGGAGTGA	1739	2760	TCATTTGCTTGAGCAACACAGGCTGTGGCTGTGTACTGATCCAGCAATATCTGCAAGG	2819
QY	1740	AGCGCTTGTGTTAGCAGCAGGACCTGCTATTCGGTGTGTGTGGAACACAGGCTGCTCA	1799	2820	GAAATGCATAGAGGTTTCTATAAGGACCAAGTGAAGATGCTTCCGCAAGCCCTACAGG	2879
Db	1740	AGCGCTTGTGTTAGCAGCAGGACCTGCTATTCGGTGTGTGTGGAACACAGGCTGCTCA	1799	2820	GAAATGCATAGAGGTTTCTATAAGGACCAAGTGAAGATGCTTCCGCAAGCCCTACAGG	2879
QY	1800	GTGTATCTGTGGCGCTGCAATGATGAACAGGAAGAAAGTTAAATCAGAAATGTTT	1859	2880	AAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGGACACAGCATACAA	2939
Db	1800	GTGTATCTGTGGCGCTGCAATGATGAACAGGAAGAAAGTTAAATCAGAAATGTTT	1859	2880	AAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGGACACAGCATACAA	2939
QY	1860	TTCCAAAGAACTCTTGACCATGACAGATGTGACACAGACACAGATTTGTTACAGCTGCAC	1919	2940	CTGGTCTTTTCAATTCATCTCCAGCTTGCCCAATGCAACGGCCACAGTAAATTCATCA	2999
Db	1860	TTCCAAAGAACTCTTGACCATGACAGATGTGACACAGACACAGATTTGTTACAGCTGCAC	1919	2940	CTGGTCTTTTCAATTCATCTCCAGCTTGCCCAATGCAACGGCCACAGTAAATTCATCA	2999
QY	1920	AGCCAAACCAATGATCTGCGCACTGGTGAATGACATTTGTCCTCCAGGAAACACAGCTG	1979	3000	GAGCATCTGTGAGAAAGTGTGAGAACCTTGACCAAGGCAAGCATGCGAGACCTGCATATC	3059
Db	1920	AGCCAAACCAATGATCTGCGCACTGGTGAATGACATTTGTCCTCCAGGAAACACAGCTG	1979	3000	GAGCATCTGTGAGAAAGTGTGAGAACCTTGACCAAGGCAAGCATGCGAGACCTGCATATC	3059
QY	1980	CTCAGAGGCGAGATCTCCATTTTATGATGAGAAATGTCCTCCAGGAAACCCCTATGTA	2039	3060	TGGCTTTTACCGGTGATCCCAATGAGGGAATGTGAGCCATGCAAGTGCATATGCGCA	3119
Db	1980	CTCAGAGGCGAGATCTCCATTTTATGATGAGAAATGTCCTCCAGGAAACCCCTATGTA	2039	3060	TGGCTTTTACCGGTGATCCCAATGAGGGAATGTGAGCCATGCAAGTGCATATGCGCA	3119
QY	2040	CTACTGTAAACAAGAACACAGCTGCGAGGCTGTGCTTGGACCAAGATGCGGAGTGGGA	2099	3120	CGCGTCTCTGTGGAACCAACACCGGCAAGTGTCTTGACACCAACGAGGCGTCAGGG	3179
Db	2040	CTACTGTAAACAAGAACACAGCTGCGAGGCTGTGCTTGGACCAAGATGCGGAGTGGGA	2099	3120	CGCGTCTCTGTGGAACCAACACCGGCAAGTGTCTTGACACCAACGAGGCGTCAGGG	3179
QY	2100	GCCCCGGAATCAGAGTGCAATGCTGCTGCGGAAATATCTGCGCATTTGGCTGGCATTT	2159	3180	GGAAGAGTGCAGCTATGTGAGGATAGAAATCGATACCAAGGAAACCTCTCAGAGGAAC	3239
Db	2100	GCCCCGGAATCAGAGTGCAATGCTGCTGCGGAAATATCTGCGCATTTGGCTGGCATTT	2159	3180	GGAAGAGTGCAGCTATGTGAGGATAGAAATCGATACCAAGGAAACCTCTCAGAGGAAC	3239
QY	2160	GGTTGGAATCTATGTTGAAATTTACTGTCAGAGGAAATTTGACAAATGCTAAAT	2219	3240	ATGTTATTTACTCTTCTTATGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGA	3299
Db	2160	GGTTGGAATCTATGTTGAAATTTACTGTCAGAGGAAATTTGACAAATGCTAAAT	2219	3240	ATGTTATTTACTCTTCTTATGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGA	3299
QY	2220	GTTCGTAGGAACCAATGCTTGTGCTTCTTCAACCCAGGAAGGATGAAAT	2279	3300	TCGCTATTTACACAGCTATCAATTTTGGGCTACTCTGACGAAACAAACAGGATTTGGA	3359
Db	2220	GTTCGTAGGAACCAATGCTTGTGCTTCTTCAACCCAGGAAGGATGAAAT	2279	3300	TCGCTATTTACACAGCTATCAATTTTGGGCTACTCTGACGAAACAAACAGGATTTGGA	3359
				3360	CATGTTTCATCAATGCTTCCAGGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTC	3419

Db 3360 CATGTCATCATGCTCCAGAAATTTCAACTCACAATCACTGGCTGGCCAGTTCTC 3419  
QY 3420 AGCTGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAATTAAGAGTA 3479  
Db 3420 AGCTGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAAACCAATTAAGAGTA 3479  
QY 3480 CAAAGATAGTTTCTTAATGAGAGATTTGATTTTGGCAACCCCAAAATATCACTTTCT 3539  
Db 3480 CAAAGATAGTTTCTTAATGAGAGATTTGATTTTGGCAACCCCAAAATATCACTTTCT 3539  
QY 3540 TGTATTATGTCAGTAATTTCACTGGCCCAATCAAAATTCAGAT 3581  
Db 3540 TGTATTATGTCAGTAATTTCACTGGCCCAATCAAAATTCAGAT 3581  
RESULT 7  
ADD70997  
ID ADD70997 standard; DNA; 3597 BP.  
XX AC ADD70997;  
XX DT 15-JAN-2004 (first entry)  
XX Human attractin gene SEQ ID NO:1.  
XX DE liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;  
XX KW cytosstatic; gene therapy; human; gene; ds.  
XX KW Homo sapiens.  
XX OS WO2003061564-A2.  
XX PN 31-JUL-2003.  
XX PD 20-DEC-2002; 2002WO-US040718.  
XX PF 21-DEC-2001; 2001US-034181SP.  
XX PR 31-DEC-2001; 2001US-034181SP.  
XX XX (GENE-) GENE LOGIC INC.  
XX PA (LGBI-) LG BIOMEDICAL INST.  
XX PI Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;  
XX WPI; 2003-663343/62.  
XX  
XX Diagnosing liver cancer cells, useful for treating liver cancer  
XX associated with chronic hepatitis or cirrhosis comprises detecting the  
XX level of expression in a tissue sample of one or more genes associated  
XX with cancerous liver tissues.  
XX  
XX Claim 1; SEQ ID NO 1; 176pp; English.  
XX  
XX The present invention describes a method for diagnosing liver cancer  
XX cells comprising detecting the level of expression in a tissue sample of  
XX one or more genes given in the specification (see ADD70997 to ADD71105),  
XX where differential expression of the genes is indicative of liver cancer.  
XX Also described: (1) detecting the progression of liver cancer in a  
XX patient; (2) monitoring the treatment of a patient with liver cancer; (3)  
XX treating a patient with liver cancer; (4) typing a liver disease in a  
XX patient; (5) detecting the presence or progression of liver cancer in a  
XX patient with chronic hepatitis or cirrhosis; (6) differentiating liver  
XX cancer related to chronic hepatitis from liver cancer related to  
XX cirrhosis; (7) screening for an agent capable of modulating the onset or  
XX progression of liver cancer; (8) a composition comprising at least two  
XX oligonucleotides comprising a sequence that specifically hybridizes to  
XX any of the genes; (9) a solid support comprising the at least two  
XX oligonucleotides; (10) a computer system comprising a database containing  
XX information identifying the level in liver tissue of a set of genes; (11)  
XX a method for using the computer system to present information identifying  
XX the expression level in tissue or cell of any of the genes; and (12) a  
XX therapeutic agent for slowing or halting the progression of liver cancer.

CC The methods are useful for treating liver cancer associated with chronic  
CC hepatitis or cirrhosis. The present sequence represents a specifically  
CC claimed human gene sequence which is used in the exemplification of the  
CC present invention.

XX SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Query Match 87.4%; Score 3556.4; DB 9; Length 3597;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 3579; Conservative 0; Mismatches 1; Indels 2; Gaps 2;  
QY 1 ATGTGGCCGAGCGGGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGGGCGACGGCA 60  
Db 1 ATGTGGCCGAGCGGGGCACTGAGGCAAGGCTGAGGAGGAGGAGCGGGCGACGGCA 60  
QY 61 GCGCTCGCGGCGAGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
Db 61 GCGCTCGCGGCGAGGCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 120  
QY 121 ACCGCGCAGTGGCTGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
Db 121 ACCGCGCAGTGGCTGTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 180  
QY 181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGTGGACCTGGAATATTAATAC 240  
Db 181 CGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGTGGACCTGGAATATTAATAC 240  
QY 241 AAAACGAAGTCACGCTGGCTCAATGAAGGACAGCAAAATAGAATAATGAGACTTCGTTTC 300  
Db 241 AAAACGAAGTCACGCTGGCTCAATGAAGGACAGCAAAATAGAATAATGAGACTTCGTTTC 300  
QY 301 AATCATTTTGTACAGAGTGTAGTGGACCAATTTATATGTTATGATGGGAGCTCAATT 360  
Db 301 AATCATTTTGTACAGAGTGTAGTGGACCAATTTATATGTTATGATGGGAGCTCAATT 360  
QY 361 TATGACCGCTAGTTGTGCAATTTAGTGGCGCTCAATTTCTGAGAGAGATGGCAATGAG 420  
Db 361 TATGACCGCTAGTTGTGCAATTTAGTGGCGCTCAATTTCTGAGAGAGATGGCAATGAG 420  
QY 421 ACTGTCCTGAGGTTGTTGCCACATCAGGTTATGCTGCTGCAATTTTGTAGTGAATGCT 480  
Db 421 ACTGTCCTGAGGTTGTTGCCACATCAGGTTATGCTGCTGCAATTTTGTAGTGAATGCT 480  
QY 481 GCTTATATTTGACTGGATTTAAATATTAATTAATTAATTAATTAATTAATTAATTAAT 540  
Db 481 GCTTATATTTGACTGGATTTAAATATTAATTAATTAATTAATTAATTAATTAATTAAT 540  
QY 541 TCAGGCGGAGAGAGTGTAGATCAGTAATAGCAGCGATCTGTTGAATGTGAATGTCT 600  
Db 541 TCAGGCGGAGAGAGTGTAGATCAGTAATAGCAGCGATCTGTTGAATGTGAATGTCT 600  
QY 601 GAAACTCGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTCT 660  
Db 601 GAAACTCGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTCT 660  
QY 661 CATCGAGCACTGCAATTTCAAGTGTGTAGAGATGCTCCTGCTCTCAGACTGGCAG 720  
Db 661 CATCGAGCACTGCAATTTCAAGTGTGTAGAGATGCTCCTGCTCTCAGACTGGCAG 720  
QY 721 GGTCTCGATGTTGATGTTCTGTTACAGCTAACAGTCAATTTTGGACTCGAGAGAAAT 780  
Db 721 GGTCTCGATGTTGATGTTCTGTTACAGCTAACAGTCAATTTTGGACTCGAGAGAAAT 780  
QY 781 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAACATTAATG 840  
Db 781 TCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAACATTAATG 840  
QY 841 GTTGTGGAGGATATATGTTCAACCACTCAGATTATTAACATGGTTCAGCGTATGACCT 900  
Db 841 GTTGTGGAGGATATATGTTCAACCACTCAGATTATTAACATGGTTCAGCGTATGACCT 900  
QY 901 GCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGTGAACATGTTGTTTAGATATG 960



D <sub>b</sub>	901	GCTTCTAGGGAGTGGCTTCCACTAAACCGTTCGTGTGAACAATGTGGTTGTTAGATATGGT	960
Q <sub>y</sub>	961	CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT	1020
D <sub>b</sub>	961	CATTCTTTGGCATTATACAAGGATAAAATTTACATGTATGGAGGAAAAATTGATTCAACT	1020
Q <sub>y</sub>	1021	GGGAATGTGACCAATGATGTGAGAGTTTTTTCACATTCATTAATGAGTCAATGGGTGTGTTG	1080
D <sub>b</sub>	1021	GGGAATGTGACCAATGATGTGAGAGTTTTTTCACATTCATTAATGAGTCAATGGGTGTGTTG	1080
Q <sub>y</sub>	1081	ACCCTTAAGGCAAAAGGAGCAGTATGCAGTGGTTGGGCACCTCGCACACATTTGTTACACTG	1140
D <sub>b</sub>	1081	ACCCTTAAGGCAAAAGGAGCAGTATGCAGTGGTTGGGCACCTCGCACACATTTGTTACACTG	1140
Q <sub>y</sub>	1141	AAGAAATGGCCGAGTGGTCACTCTGGTCACTCTTTTGGTCACTGGCCCTCTCTATGAGATATATA	1200
D <sub>b</sub>	1141	AAGAAATGGCCGAGTGGTCACTCTGGTCACTCTTTTGGTCACTGGCCCTCTCTATGAGATATATA	1200
Q <sub>y</sub>	1201	AGCAATGTGCAGGAATATGATTTTGGATAAGAAACAATGGAGTATATTACACACCCAGGGT	1260
D <sub>b</sub>	1201	AGCAATGTGCAGGAATATGATTTTGGATAAGAAACAATGGAGTATATTACACACCCAGGGT	1260
Q <sub>y</sub>	1261	GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCGGGGCCCTA	1320
D <sub>b</sub>	1261	GCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACCGGGGCCCTA	1320
Q <sub>y</sub>	1321	TACGTTTCATGGTGGCTACAAAGGCTTCAGTGGCCCAATAGTACCGGCTTCAGAGTATCTC	1380
D <sub>b</sub>	1321	TACGTTTCATGGTGGCTACAAAGGCTTCAGTGGCCCAATAGTACCGGCTTCAGAGTATCTC	1380
Q <sub>y</sub>	1381	TACCGATATGATGTGGATACCAGATGTGCACCAATCTCTTAAGGACAGCCGATTTTCCCGT	1440
D <sub>b</sub>	1381	TACCGATATGATGTGGATACCAGATGTGCACCAATCTCTTAAGGACAGCCGATTTTCCCGT	1440
Q <sub>y</sub>	1441	TACTTGCACACAGCTGTGATAGTGTGAGTGGAAACCAATGTGTGTGTGGGGGAAACACACAC	1500
D <sub>b</sub>	1441	TACTTGCACACAGCTGTGATAGTGTGAGTGGAAACCAATGTGTGTGTGGGGGAAACACACAC	1500
Q <sub>y</sub>	1501	AATGACACATCTATGAGCCATGCGCCCAATGCTCTCTTCAGATTTTCATGGCCCTATGAC	1560
D <sub>b</sub>	1501	AATGACACATCTATGAGCCATGCGCCCAATGCTCTCTTCAGATTTTCATGGCCCTATGAC	1560
Q <sub>y</sub>	1561	ATTGCTGTGACCGCTGGTCACTGTCTCCAGACCTGAT-TCCACCATGATGTCAACAGA	1620
D <sub>b</sub>	1561	ATTGCTGTGACCGCTGGTCACTGTCTCCAGACCTGAT-TCCACCATGATGTCAACAGA	1619
Q <sub>y</sub>	1621	TTTGGCCATT-CAGCAGTCTTACACAACAGCACCATGTATGTCTTCGGTGGTTTCCATAG	1679
D <sub>b</sub>	1620	TTTGGCCATTTCACAGCAGTCTTACACAACAGCACCATGTATGTCTTCGGTGGTTTCCATAG	1679
Q <sub>y</sub>	1680	TCTCTCTCTCAGCGACATCCTGTGTATTTCACCTGGAAACAGTGTGATCGCATCGGAGTGA	1739
D <sub>b</sub>	1680	TCTCTCTCTCAGCGACATCCTGTGTATTTCACCTCGGAACAGTGTGATCGCATCGGAGTGA	1739
Q <sub>y</sub>	1740	AGCGGCTTTTATGACAGCAGACCTGTGTATTCGGTGTGTGTGAAACACAGGTCGTCTCA	1799
D <sub>b</sub>	1740	AGCGGCTTTTATGACAGCAGACCTGTGTATTCGGTGTGTGTGAAACACAGGTCGTCTCA	1799
Q <sub>y</sub>	1800	GTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAGAAAGTTAAATCAGAATGTTT	1859
D <sub>b</sub>	1800	GTGTATCTCGTGGCGCTGGCAACTGATGAACAAGAGAAAGTTAAATCAGAATGTTT	1859
Q <sub>y</sub>	1860	TTCCAAAAGAACTCTTGAACCATGACAGATGTGACCAAGCAACAGATTTGTTACAGTGCAC	1919
D <sub>b</sub>	1860	TTCCAAAAGAACTCTTGAACCATGACAGATGTGACCAAGCAACAGATTTGTTACAGTGCAC	1919
Q <sub>y</sub>	1920	AGCCACACCAATGACTGCCACTGGTCCATGACCATTTGTGTCCCGAGGAACCAACAGCTG	1979
D <sub>b</sub>	1920	AGCCACACCAATGACTGCCACTGGTCCATGACCATTTGTGTCCCGAGGAACCAACAGCTG	1979
Q <sub>y</sub>	1980	CTCAGAAGGCCAGATCTCCATTTTATAGGTATGAGAAATTGCCCCAAGATTAACCTATGTA	2039
D <sub>b</sub>	1980	CTCAGAAGGCCAGATCTCCATTTTATAGGTATGAGAAATTGCCCCAAGATTAACCTATGTA	2039

QY	2040	CTACTGTAAACAAGAGACCAAGCTGCAGAGAGCTGTGCCCTTGACACAGAACTGCCCAGTGGGA	2099
DB	2040	CTACTGTAAACAAGAGACCAAGCTGCAGAGAGCTGTGCCCTTGACACAGAACTGCCCAGTGGGA	2099
QY	2100	GCCCGGGAATCAGGAGTGCATTGCCCTGCCCGAAATATCTGTGGCAATGGCTGGCAATTT	2159
DB	2100	GCCCGGGAATCAGGAGTGCATTGCCCTGCCCGAAATATCTGTGGCAATGGCTGGCAATTT	2159
QY	2160	GGTTGGAAAATCTCATGTTTGGAAAATTAATCTACGCCAAGAGAAATTTATGACAAATGCTAAATTT	2219
DB	2160	GGTTGGAAAATCTCATGTTTGGAAAATTAATCTACGCCAAGAGAAATTTATGACAAATGCTAAATTT	2219
QY	2220	GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTCAACACCCAGAGAAGGTAGAAATTT	2279
DB	2220	GTTCTGTAGGAACCACAATGCCCTTTTGGCTTCTCTTCAACACCCAGAGAAGGTAGAAATTT	2279
QY	2280	TGTCCTTTAAGCAGCTGGCAATATGTCAGTGCATCTTCAGAGCATGTCCAAAGTCACTTTAAAC	2339
DB	2280	TGTCCTTTAAGCAGCTGGCAATATGTCAGTGCATCTTCAGAGCATGTCCAAAGTCACTTTAAAC	2339
QY	2340	CCCATGGGTGGCCCTTCGGAAAGTCAATGTGTCTCTACTGTGTCTGGGAAGATATGTCCCC	2399
DB	2340	CCCATGGGTGGCCCTTCGGAAAGTCAATGTGTCTCTACTGTGTCTGGGAAGATATGTCCCC	2399
QY	2400	ATTTTACAANTAGTTTACTACAGTGGATGCCGTCTGAGCCAGTCAGTGCATGCTGGATCTCTGG	2459
DB	2400	ATTTTACAANTAGTTTACTACAGTGGATGCCGTCTGAGCCAGTCAGTGCATGCTGGATCTCTGG	2459
QY	2460	AAATTTTATCAGAACCCAGTACTCGGGGACTCAAGGCTGCACCTCTGCATCAACCCACTCAA	2519
DB	2460	AAATTTTATCAGAACCCAGTACTCGGGGACTCAAGGCTGCACCTCTGCATCAACCCACTCAA	2519
QY	2520	TGGTAGTGTCTGTGAAAGGCTGCACCAACCAAGTGTCTAAGCAGTGGCCGACACCATGTGC	2579
DB	2520	TGGTAGTGTCTGTGAAAGGCTGCACCAACCAAGTGTCTAAGCAGTGGCCGACACCATGTGC	2579
QY	2580	CTTGAGGACAGCATGTGGAGATTGCACACGCGGAGCTCTGAGTGCATGTGGTGACAGAA	2639
DB	2580	CTTGAGGACAGCATGTGGAGATTGCACACGCGGAGCTCTGAGTGCATGTGGTGACAGAA	2639
QY	2640	CATGAAGCAGTGTGACACTCCAATGCCATGTGGCCTCTCTCCCTTTTGGCCAGTGTAT	2699
DB	2640	CATGAAGCAGTGTGACACTCCAATGCCATGTGGCCTCTCTCCCTTTTGGCCAGTGTAT	2699
QY	2700	GGAAATGTTATACGATGAGCACCTGCCCCCTCGAAAATTTGTTTCAGGCTACTGTACCTGTAG	2759
DB	2700	GGAAATGTTATACGATGAGCACCTGCCCCCTCGAAAATTTGTTTCAGGCTACTGTACCTGTAG	2759
QY	2760	TCATTGCTTCGAGCAACCAAGCTGTGCTGGTGTACTGATGCCAGCAATACTGCGAAGG	2819
DB	2760	TCATTGCTTCGAGCAACCAAGCTGTGCTGGTGTACTGATGCCAGCAATACTGCGAAGG	2819
QY	2820	GAATGATGAGAGGTTTCTTATAAAGACCAAGTGAAGATGCCCTTCGAAGCCCTTACAGG	2879
DB	2820	GAATGATGAGAGGTTTCTTATAAAGACCAAGTGAAGATGCCCTTCGAAGCCCTTACAGG	2879
QY	2880	AAATTTCTATCCACAGCCCTGTCTCAATTCAGCATGTGTAGAGACAGCAGATACAA	2939
DB	2880	AAATTTCTATCCACAGCCCTGTCTCAATTCAGCATGTGTAGAGACAGCAGATACAA	2939
QY	2940	CTGGTCTTTTCAATTCATGTCCAGCTTGCCAAATGCAACGGCCACAGTAAATGCAATCA	2999
DB	2940	CTGGTCTTTTCAATTCATGTCCAGCTTGCCAAATGCAACGGCCACAGTAAATGCAATCA	2999
QY	3000	GAGCATCTGTGAGAAGTGTGAGAACCTTGACACAGGCAAGCACTTGGAGACCTTGCATATC	3059
DB	3000	GAGCATCTGTGAGAAGTGTGAGAACCTTGACACAGGCAAGCACTTGGAGACCTTGCATATC	3059
QY	3060	TGGCTTCTTACGGTATCCCAACCAATGGAGGGAATGTTCAGCCATGCAAGTGCATGGGCA	3119
DB	3060	TGGCTTCTTACGGTATCCCAACCAATGGAGGGAATGTTCAGCCATGCAAGTGCATGGGCA	3119





Db 954 AGTTCCTGTACAGTAAACAGTCAITTTTGGACTCGAGAGGATATTTAACTTAAAGCT 1013  
Qy 795 CCCGAGAGCATCTCATTAAGCTGTGTCAATGGAACATTTATGTGGTGTGTGGAGGATA 854  
Db 1014 CCCGAGAGCATCTCATTAAGCTGTGTCAATGGAACATTTATGTGGTGTGTGGAGGATA 1073  
Qy 855 TATGTTCAACCACTCAAGTATTAACATGTTTCTAGGATATGACTTGTCTTCTAGGAGTG 914  
Db 1074 TATGTTCAACCACTCAAGTATTAACATGTTTCTAGGATATGACTTGTCTTCTAGGAGTG 1133  
Qy 915 GCTTCCACTAAACCGTCTCTGTGAACATGTTGTTGTAGATATGTTGTCATTTCTTGGCAT 974  
Db 1134 GCTTCCACTAAACCGTCTCTGTGAACATGTTGTTGTAGATATGTTGTCATTTCTTGGCAT 1193  
Qy 975 ATCAAGGATATAAATTTACATGATGAGGAGGAAATTTGATTCAACTGGGAATGTGACCAA 1034  
Db 1194 ATCAAGGATATAAATTTACATGATGAGGAGGAAATTTGATTCAACTGGGAATGTGACCAA 1253  
Qy 1035 TGAGTTGAGAGTTTTCACATTTCAATATGATGATGATGATGATGATGATGATGATGATGAT 1094  
Db 1254 TGAGTTGAGAGTTTTCACATTTCAATATGATGATGATGATGATGATGATGATGATGATGAT 1313  
Qy 1095 GGAGCATATGAGTGTGGGCACTCTGCACATTTGTTTACACTGAAGATGGCGAGT 1154  
Db 1314 GGAGCATATGAGTGTGGGCACTCTGCACATTTGTTTACACTGAAGATGGCGAGT 1373  
Qy 1155 GGTCACTGCTGTCATCTTGTGTCATCTCTCTATGATATATAAGCAATGTGCAGGA 1214  
Db 1374 GGTCACTGCTGTCATCTTGTGTCATCTCTCTATGATATATAAGCAATGTGCAGGA 1433  
Qy 1215 ATATGATTTGGATGAAGAACATGAGGATATATTAACACACCCAGGTTGCCCTGTGCAAG 1274  
Db 1434 ATATGATTTGGATGAAGAACATGAGGATATATTAACACACCCAGGTTGCCCTGTGCAAG 1493  
Qy 1275 GGGTTACGGCCATAGCAGTGTGTACAGACATAGGACAGGGCCCTATACGTTTCATGGTGG 1334  
Db 1494 GGGTTACGGCCATAGCAGTGTGTACAGACATAGGACAGGGCCCTATACGTTTCATGGTGG 1553  
Qy 1335 CTCAAGGCTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTAOCGATATGATGT 1394  
Db 1554 CTCAAGGCTTTCAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTAOCGATATGATGT 1613  
Qy 1395 GGTATCCAGATGTGGACATTTCTTAAGGACCGGATTTTCCGTTACTTGCACACAGC 1454  
Db 1614 GGTATCCAGATGTGGACATTTCTTAAGGACCGGATTTTCCGTTACTTGCACACAGC 1673  
Qy 1455 TGTGATAGTGAAGGAAACCATGCTGTGTGTTTGGGGGAAACACACAAATGACATCTAT 1514  
Db 1674 TGTGATAGTGAAGGAAACCATGCTGTGTGTTTGGGGGAAACACACAAATGACATCTAT 1733  
Qy 1515 GAGCCATGGCCCAATGCTTCTCTCAGATTTTCAGGCTATGACATGCTGTGACCG 1574  
Db 1734 GAGCCATGGCCCAATGCTTCTCTCAGATTTTCAGGCTATGACATGCTGTGACCG 1793  
Qy 1575 CTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634  
Db 1794 CTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853  
Qy 1635 AGTCTTACAAACAGCACCATATGATGTTTGGTGTGTTCAATAGTCTCGTCTCAGCCA 1694  
Db 1854 AGTCTTACAAACAGCACCATATGATGTTTGGTGTGTTCAATAGTCTCGTCTCAGCCA 1913  
Qy 1695 CATCTGGTATTCACCTCGGAAACAGTGTGATGCGCATCGAGTGAAGCGCTGTGTTTACG 1754  
Db 1914 CATCTGGTATTCACCTCGGAAACAGTGTGATGCGCATCGAGTGAAGCGCTGTGTTTACG 1973  
Qy 1755 AGCAGGACCTGGTATTCGGTGTGTGGAACAACAGGGTGTCTCAGTGTATCTCGTGGGC 1814  
Db 1974 AGCAGGACCTGGTATTCGGTGTGTGGAACAACAGGGTGTCTCAGTGTATCTCGTGGGC 2033  
Qy 1815 GCTGGCACTGATGAACAAGAAAGTTTAAATCAGAAATGTTTTCAAAAGAACTCT 1874

2034 GCTGGCACTGATGAACAAGAAAGTTTAAATCAGAAATGTTTTCAAAAGAACTCT 2093  
1875 TGACCATGACAGATGTGACACAGCACACAGATTGTTTACAGTGTGACACCAACCAATGA 1934  
2094 TGACCATGACAGATGTGACACAGCACACAGATTGTTTACAGTGTGACACCAACCAATGA 2153  
1935 CTGCCACTGTGCAATGACCACTTGTGCTCCAGGAAACACAGCTGTCTCAGAAAGCCAGAT 1994  
2154 CTGCCACTGTGCAATGACCACTTGTGCTCCAGGAAACACAGCTGTCTCAGAAAGCCAGAT 2213  
1995 CTCCATTTTATAGTATGAGATTCGCCCAAGGATAACCTCTATGTACTACTGTAAACAGAA 2054  
2214 CTCCATTTTATAGTATGAGATTCGCCCAAGGATAACCTCTATGTACTACTGTAAACAGAA 2273  
2055 GACAGCTGTGACAGGAGTGTGCCCTGGACACAGAACTGCGAGTGGAGGCCCGGAATCAGGA 2114  
2274 GACAGCTGTGACAGGAGTGTGCCCTGGACACAGAACTGCGAGTGGAGGCCCGGAATCAGGA 2333  
2115 GTGCATTTGCCCTGCCGGAATATCTGTGCAATTTGGCTGGCAATTTGGTTGGAACTCATG 2174  
2334 GTGCATTTGCCCTGCCGGAATATCTGTGCAATTTGGCTGGCAATTTGGTTGGAACTCATG 2393  
2175 TTTGAAATTTACTACTGCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACCA 2234  
2394 TTTGAAATTTACTACTGCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACCA 2453  
2235 CAATGCCCTTTTGGCTTCTCTTCAACCCAGAGAGTATGATGCTTCTTCTTAAAGAGCT 2294  
2454 CAATGCCCTTTTGGCTTCTCTTCAACCCAGAGAGTATGATGCTTCTTCTTAAAGAGCT 2513  
2295 GCGAATATGAGTCACTCTCAGACATGTCACAGCTCACTTAAACCCCAATGGGTGCGCCT 2354  
2514 GCGAATATGAGTCACTCTCAGACATGTCACAGCTCACTTAAACCCCAATGGGTGCGCCT 2573  
2355 TCGAAGATCAATGTGCTTCTACTGCTGGGAAAGATATGTGCCCATTTTCAAAATAGTTT 2414  
2574 TCGAAGATCAATGTGCTTCTACTGCTGGGAAAGATATGTGCCCATTTTCAAAATAGTTT 2633  
2415 ACTACAGTGGATGCTGAGCCAGTGTGCTGAGTCTGTGGATTTCTGTGGAATTTTATCAGAAC 2474  
2634 ACTACAGTGGATGCTGAGCCAGTGTGCTGAGTCTGTGGATTTCTGTGGAATTTTATCAGAAC 2693  
2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCTCAACCCACTCAATGCTGTCTGTGA 2534  
2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCTCAACCCACTCAATGCTGTCTGTGA 2753  
2535 AAGGCTTGCACCAACACAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAG 2594  
2754 AAGGCTTGCACCAACACAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAG 2813  
2595 TGGAGATTCACCAACACAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAG 2654  
2814 TGGAGATTCACCAACACAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAG 2873  
2655 GGAATCCAAATGCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2714  
2874 GGAATCCAAATGCTTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2933  
2715 GAGCACCTGCCCCCTGAAATTTCTCAGGCTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACT 2774  
2934 GAGCACCTGCCCCCTGAAATTTCTCAGGCTACTGTACTGTACTGTACTGTACTGTACTGTACTGTACT 2993  
2775 ACCAGGCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2834  
2994 ACCAGGCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3053  
2835 TTCTTATTAAGGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2894  
3054 TTCTTATTAAGGACAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3113  
2895 GCCCTTGTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCAATCA 2954  
3114 GCCCTTGTCTCAATTTCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCAATCA 3173

```
QY 2955 CTGTCAGCTTGCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAA 3014
Db 3174 CTGTCCAGCTTGCAATGCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAA 3233
QY 3015 GTGTGAGACTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAGGTCGA 3074
Db 3234 GTGTGAGAACTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTAGGTCGA 3293
QY 3075 TCCACCAATGAGGAGAAATGTGAGCCATGCAAGTGCATGAGGCAAGCGCTCTGTGCAA 3134
Db 3294 TCCACCAATGAGGAGAAATGTGAGCCATGCAAGTGCATGAGGCAAGCGCTCTGTGCAA 3353
QY 3135 CACCAACAGGCAAGTCTTCTGCACCAACCAAGGCGCTCAGGGGAGAGTGCAGCT 3194
Db 3354 CACCAACAGGCAAGTCTTCTGCACCAACCAAGGCGCTCAGGGGAGAGTGCAGCT 3413
QY 3195 ATGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGAACATGTTATATCTCT 3254
Db 3414 ATGTGAGTAGAATAATCGATACCAAGAAACCTCTCAGAGGAACATGTTATATCTCT 3473
QY 3255 TCTATTGACTATCAGTTACCTTTAGTCTATCCCAAGGAGATGTCGCTATTACACGC 3314
Db 3474 TCTATTGACTATCAGTTACCTTTAGTCTATCCCAAGGAGATGTCGCTATTACACGC 3533
QY 3315 TATCAATTTTGTGCTACTCTGACGAACAACAGGAGATTTGGACATGTTCAATCAATGC 3374
Db 3534 TATCAATTTTGTGCTACTCTGACGAACAACAGGAGATTTGGACATGTTCAATCAATGC 3593
QY 3375 CTCCAAGAAATTCACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3434
Db 3594 CTCCAAGAAATTCACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGC 3653
QY 3435 TGGAGAGAGATGCTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494
Db 3654 TGGAGAGAGATGCTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713
QY 3495 TAATGAGAGATTTGATTTTGGCAACCAACCAATATCACTTTCTTTTGTATGTCAAGTAA 3554
Db 3714 TAATGAGAGATTTGATTTTGGCAACCAACCAATATCACTTTCTTTTGTATGTCAAGTAA 3773
QY 3555 TTTCACTGCGCCCATCAAAATTCAGAT 3581
Db 3774 TTTCACTGCGCCCATCAAAATTCAGAT 3800

RESULT 9
AAZ52277
ID AAZ52277 standard; cDNA; 3819 BP.
AC AAZ52277;
DT 18-JUL-2000 (first entry)
DE Human soluble attractin-2 cDNA.
XX
XX Human; soluble attractin-2; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..3819
XX FT /*tag= a
XX FT /product= "Soluble attractin-2"
XX
XX WO200015651-A1.
XX
XX 23-MAR-2000.
XX
XX 14-SEP-1999; 99WO-US020948.
XX
```

```
PR 14-SEP-1998; 98US-0100137P.
XX
XX (DAND ) DANA FARBER CANCER INST INC.
XX
XX Duke-Cohan JS, Schlosman SF;
XX
XX WPI; 2000-271373/23.
XX P-PSDB; AAY70692.
XX
XX Isolated nucleic acids encoding human attractin polypeptides useful for
XX enhancing immune responses.
XX
XX Claim 3; Fig 12; 120pp; English.
XX
XX The patent discloses four forms of human attractin polypeptides which
XX enhance immune response by promoting macrophage and monocyte spreading in
XX the presence of T cells. These include soluble attractin-1 and -2 and
XX membrane attractin-1 and -2. These various forms of attractin are encoded
XX by alternatively spliced mRNA molecule transcribed from a single gene.
XX The present sequence is a cDNA encoding human soluble attractin-2.
XX Attractin-2 differs from attractin-1 in having a 74 amino acid insert in
XX the N-terminal. This sequence can be used to enhance immune response in
XX immunosuppressed patients such as those undergoing chemo- and radio-
XX therapy treatment for cancer or those suffering from common variable
XX immunodeficiency syndrome. The proteins may also be used to screen
XX modulators (agonists and antagonists) of immune response which may also
XX be used to regulate immune reactions. Attractin antibodies can be used to
XX inhibit immune response in transplant recipients or patients afflicted
XX with autoimmune disease
XX
XX Sequence 3819 BP; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;
XX
XX Query Match 85.9%; Score 3492.4; DB 3; Length 3819;
XX Best Local Similarity 99.3%; Pred. No. 0;
XX Matches 3508; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 48 GCGCGCAGCGCAGCGCTCGCGGCAGGAGCGCGCGGCCGCACTGTGTCAACGCGGTGCG 107
Db 270 GTCGGGCTCAGCGCAGCGCGGCAAGAAATGTGACCGGCCCTGTGTCAACGCGGTGCG 329
QY 108 CTGCACCCCTGGCACCGGCCAGTGTGTCGCCCGCGGCTGGTGGGGGAGCAATGCCA 167
Db 330 CTGCACCCCTGGCACCGGCCAGTGTGTCGCCCGCGGCTGGTGGGGGAGCAATGCCA 389
QY 168 GCACCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGACCTGG 227
Db 390 GCACCTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTTGTGACAGATGACCTGG 449
QY 228 AAATTAATAATACAAAGCAAGTGCAGTGCCTCAATTGAAGGACAGCCAAATAGATAAT 287
Db 450 AAATTAATAATACAAAGCAAGTGCAGTGCCTCAATTGAAGGACAGCCAAATAGATAAT 509
QY 288 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGACCATTTATATGTTATGA 347
Db 510 GAGACTTCGTTTCAATCAATTTTGTACAGAGTGTAGTTGGACCATTTATATGTTATGA 569
QY 348 TGGGACTCAATTTATGACCGCTAGTGTGCAATTTAGTGGCCCTCAATTGTCCTGAGAG 407
Db 570 TGGGACTCAATTTATGACCGCTAGTGTGCAATTTAGTGGCCCTCAATTGTCCTGAGAG 629
QY 408 AGATGGCAATGAGACTGTCCCTGAGGTTGTGCCACATCAGGTTATGCTTGTGTCATTT 467
Db 630 AGATGGCAATGAGACTGTCCCTGAGGTTGTGCCACATCAGGTTATGCTTGTGTCATTT 689
QY 468 TTTTAGTGATGCTGCTTATATATTTGACTGGATTTAAATTTACTTTACAGTTTGTATGTG 527
Db 690 TTTTAGTGATGCTGCTTATATATTTGACTGGATTTAAATTTACTTTACAGTTTGTATGTG 749
QY 528 TCCAAATAACTGCTCAGGCCGAGAGAGTGAAGATCAGTAATAGCAGCGATCTGTGA 587
Db 750 TCCAAATAACTGCTCAGGCCGAGAGAGTGAAGATCAGTAATAGCAGCGAACTGTGA 809
QY 588 ATGTGATGTTTCTGAAAACCTGAAAAGGTGAAGCATGTGACATTTCTCTACTGTACAGCAA 647
```

Db 810 ATGTGATGTTCTGAAACCTGGAAAGTGAAGCATGTGACATTCCTCAGTGTACAGACAA 869  
Qy 648 CTGTGGTTTCTCATCGAGGCATCTGCAATTCAGATGATGTCAGAGGATGCTCTGCTT 707  
Db 870 CTGTGGTTTCTCATCGAGGCATCTGCAATTCAGATGATGTCAGAGGATGCTCTGCTT 929  
Qy 708 CTCAGATGTCAGAGGCTCTCGATGTTGAGTTCAGTTCCTGTACAGCTAAACAGTCAATTTGGAC 767  
Db 930 CTCAGATGTCAGAGGCTCTCGATGTTGAGTTCAGTTCCTGTACAGCTAAACAGTCAATTTGGAC 989  
Qy 768 TCGAGAGGAATATTTCTAACTTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGTCAATGG 827  
Db 990 TCGAGAGGAATATTTCTAACTTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGTCAATGG 1049  
Qy 828 AAACATTTATGTGGTGTCTGGAGGATATATGTTCAACCACTCAGATTATAACATGCTTCT 887  
Db 1050 AAACATTTATGTGGTGTCTGGAGGATATATGTTCAACCACTCAGATTATAACATGCTTCT 1109  
Qy 888 AGCTATGACCTTGCTCTAGGAGTGGCTTCCACTAAACCGTTCGTGAAACAATGTGT 947  
Db 1110 AGCTATGACCTTGCTCTAGGAGTGGCTTCCACTAAACCGTTCGTGAAACAATGTGT 1169  
Qy 948 TGTTAGATATGTCATTTTGGCATATATACAGAGGATAAATTTACATGATGGAGAAA 1007  
Db 1170 TGTTAGATATGTCATTTTGGCATATATACAGAGGATAAATTTACATGATGGAGAAA 1229  
Qy 1008 AATTGATTCACCTGGGAATGTGACCAATGATGTTGAGAGTTCATGATTCATTAATCAGTCT 1067  
Db 1230 AATTGATTCACCTGGGAATGTGACCAATGATGTTGAGAGTTCATGATTCATTAATCAGTCT 1289  
Qy 1068 ATGGGCTGTGTGACCTTAAGCAAGAGGAGAGTATGCAAGTGGTGGGCACTCTGCA 1127  
Db 1290 ATGGGCTGTGTGACCTTAAGCAAGAGGAGAGTATGCAAGTGGTGGGCACTCTGCA 1349  
Qy 1128 CATTGTTACATGAGGAATGCGGAGTGTGATGCTGGTTCATCTTGGTTCAGTCCCTCT 1187  
Db 1350 CATTGTTACATGAGGAATGCGGAGTGTGATGCTGGTTCATCTTGGTTCAGTCCCTCT 1409  
Qy 1188 CTATGATATATAGCAATGTGAGGAATATGATTTGGATTAAGAACACATGGAGTATAT 1247  
Db 1410 CTATGATATATAGCAATGTGAGGAATATGATTTGGATTAAGAACACATGGAGTATAT 1469  
Qy 1248 ACACACCCAGGCTCTGTGCAAGGGGTTACGGCCATAGCAGTGTTCACACCATAG 1307  
Db 1470 ACACACCCAGGCTCTGTGCAAGGGGTTACGGCCATAGCAGTGTTCACACCATAG 1529  
Qy 1308 GACCAGGCTCTATACCTTTCATGCTGCTACAGGCTTTTCAGTGCCTAAATGACCGGCT 1367  
Db 1530 GACCAGGCTCTATACCTTTCATGCTGCTACAGGCTTTTCAGTGCCTAAATGACCGGCT 1589  
Qy 1368 TGCAGATGATCTACCCGATATAGTGGATACCCAGATGTGACCATTCCTTAAGGACAG 1427  
Db 1590 TGCAGATGATCTACCCGATATAGTGGATACCCAGATGTGACCATTCCTTAAGGACAG 1649  
Qy 1428 CCGATTTTTCGTTTACCTGACACAGCTGTGATGAGTGGAGCAACATGCTGTGTGG 1487  
Db 1650 CCGATTTTTCGTTTACCTGACACAGCTGTGATGAGTGGAGCAACATGCTGTGTGG 1709  
Qy 1488 GGGAAACACACAAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTTCAGATTT 1547  
Db 1710 GGGAAACACACAAATGACACATCTATGAGCCATGGCGCCAAATGCTTCTTCAGATTT 1769  
Qy 1548 CATTGGCTTATGACATTTGCTGTGACCGCTGGTTCAGTGTCTCCAGACCTGATCTCCACCA 1607  
Db 1770 CATTGGCTTATGACATTTGCTGTGACCGCTGGTTCAGTGTCTCCAGACCTGATCTCCACCA 1829  
Qy 1608 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAAACAGACCATGATGTGTGG 1667  
Db 1830 TGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAAACAGACCATGATGTGTGG 1889  
Qy 1668 TGGTTTCAAATGATCTCTCTCAGCGACATCTGTTGATTTACCTCCGAAACAGTGTGATGC 1727

1890 TGGTTTCAAATGATCTCTCTCTCAGCGACATCTCTGTTATTCACCTCGGAACAGTGTGATGC 1949  
Qy 1728 GCATCGGAGTGAAGCCCTTGTATTAGCAGCAGGACCTGTTATTCGGTGTGTGTGGAACAC 1787  
Db 1950 GCATCGGAGTGAAGCCCTTGTATTAGCAGCAGGACCTGTTATTCGGTGTGTGTGGAACAC 2009  
Qy 1788 AGGTCGCTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAAGTTTAAA 1847  
Db 2010 AGGTCGCTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAGAAAGTTTAAA 2069  
Qy 1848 ATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACACGACACAGATTG 1907  
Db 2070 ATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACACGACACAGATTG 2129  
Qy 1908 TTACAGCTTGACAGCAACCAATGACTGCTGCACTGGTGAACATGATGTTGTCCCGAG 1967  
Db 2130 TTACAGCTTGACAGCAACCAATGACTGCTGCACTGGTGAACATGATGTTGTCCCGAG 2189  
Qy 1968 GAAACACAGCTGCTCAGAGGCCAGATCTCAATTTTTAGGTATGAGAAATGCCCCAAGGA 2027  
Db 2190 GAAACACAGCTGCTCAGAGGCCAGATCTCAATTTTTAGGTATGAGAAATGCCCCAAGGA 2249  
Qy 2028 TAAACCTATGTACTACTGTAAACAAGAACAGCTGCAAGGAGCTGTGCTCGAACACAGAA 2087  
Db 2250 TAAACCTATGTACTACTGTAAACAAGAACAGCTGCAAGGAGCTGTGCTCGAACACAGAA 2309  
Qy 2088 CTGCCAGTGGAGCCCGGAATCAGGAGTGCAATTTGCCCTGCCCCGAAATATCTGTGGCAT 2147  
Db 2310 CTGCCAGTGGAGCCCGGAATCAGGAGTGCAATTTGCCCTGCCCCGAAATATCTGTGGCAT 2369  
Qy 2148 TGGCTGGCAATTTGGTGGAACTCATGTTGAAATTTACTACTGCCAAGAGAAATTAAGA 2207  
Db 2370 TGGCTGGCAATTTGGTGGAACTCATGTTGAAATTTACTACTGCCAAGAGAAATTAAGA 2429  
Qy 2208 CAATGCTAAATTTGTTCTGTAGGAACCAATGCCCTTTTGGCTTCTCTTCAACCCAGAA 2267  
Db 2430 CAATGCTAAATTTGTTCTGTAGGAACCAATGCCCTTTTGGCTTCTCTTCAACCCAGAA 2489  
Qy 2268 GAAAGTAGAATTTGCTTAAAGCAGCTGCGAATTAATGACATCATCTCAGAGCATGTCCAA 2327  
Db 2490 GAAAGTAGAATTTGCTTAAAGCAGCTGCGAATTAATGACATCATCTCAGAGCATGTCCAA 2549  
Qy 2328 GCTCACCTTAAACCCCATCGGTCGGCTTGGAAAGATCAATGTGTCTCTACTGGTCTGGGA 2387  
Db 2550 GCTCACCTTAAACCCCATCGGTCGGCTTGGAAAGATCAATGTGTCTCTACTGGTCTGGGA 2609  
Qy 2388 AGATATGTCCTCAATTAACAATAGTTTACTACAGTGGATGCGCTGAGGCCAGTGTATGC 2447  
Db 2610 AGATATGTCCTCAATTAACAATAGTTTACTACAGTGGATGCGCTGAGGCCAGTGTATGC 2669  
Qy 2448 TGGATTCTGTGGAAATTTTATCAGAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2507  
Db 2670 TGGATTCTGTGGAAATTTTATCAGAAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCAT 2729  
Qy 2508 CAACCCACTCAATGCTAGTGTGTGTGAAGGCCCTGCAAAACCAAGTGTCTAAGCAGTGCCTG 2567  
Db 2730 CAACCCACTCAATGCTAGTGTGTGTGAAGGCCCTGCAAAACCAAGTGTCTAAGCAGTGCCTG 2789  
Qy 2568 GACACCATGTCCTTGAGGACAGCATGTGGAGATTGACACAGCGGAGCTCTGAGTGCAT 2627  
Db 2790 GACACCATGTCCTTGAGGACAGCATGTGGAGATTGACACAGCGGAGCTCTGAGTGCAT 2849  
Qy 2628 GTGTGACAGCAACATGAGCAGTGTGTGGAATCCAAATGCCCTATGTGCCCTCTCTTCCCTTT 2687  
Db 2850 GTGTGACAGCAACATGAGCAGTGTGTGGAATCCAAATGCCCTATGTGCCCTCTCTTCCCTTT 2909  
Qy 2688 TGGCCAGTGTATGGAATGGTATACATGAGCAGCTCTGCCCTCTGAAATTTGTTCCAGCTA 2747  
Db 2910 TGGCCAGTGTATGGAATGGTATACATGAGCAGCTCTGCCCTCTGAAATTTGTTCCAGCTA 2969  
Qy 2748 CTGTACCTGTAGTCAATTTGAGGACCAACAGGCTGTGGCTGTGTACTGATCCAGCAA 2807  
Db 2970 CTGTACCTGTAGTCAATTTGAGGACCAACAGGCTGTGGCTGTGTACTGATCCAGCAA 3029



255 GTGGCTCATTTGAGGACAGCCAAATAGAAATAGACACTTCGTTTCAATCATTTTCTAC 314  
Db  
546 ATGGCTCATTTGAGGACAGCCAAATAGAAATAGACACTTCGTTTCAATCATTTTCTAC 605  
Qy  
315 AGAGTGTAGTTGGACACATTTATATGTATATATGATGGGACTCAATTTATGACCGTAGT 374  
Db  
606 TGAATGTAGCTGGACACATTTATATGTATATATGATGGGACTCAATTTAAGCACCTCTGAT 665  
Qy  
375 TGCTGCATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGCTCCTGAGT 434  
Db  
666 TGCTGCCTTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGCTCCTGAGT 725  
Qy  
435 TGTTGCCACATCAGTTATGCTTGTGCTATTTTATGATGCTGCTTATATATTTGAC 494  
Db  
726 CACTGTCACTTCAGTTATGCACTACTGCAATTTTTCAGTGTGCTTATATATCTGAC 785  
Qy  
495 TGAATTTAATTTACTTTACAGTTTGTATGTGTCCAAATTAATCTGCTAGGCCGAGAGA 554  
Db  
786 TGAATTTAATTTACTTTACAGTTTGTATGTGTCCAAATTAATTTGCTCAGGCCGAGAGA 845  
Qy  
555 GTGTAAGATCAGTAAATAGCAGCGATCTGTTGAATGTGAATGTTCTGAAACTTGAAGG 614  
Db  
846 ATGTAAGAGCATTAACAGCAGCAGCTGTTGAGTGTGAATGTTCTGAAACTTGAAGG 905  
Qy  
615 TGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCATCTG 674  
Db  
906 GGAGTCTGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCATCTG 965  
Qy  
675 CAATTTAAGTGTGTGAGAGTGTCTCTGCTCTCAGACTGCGAGGTCCTGATGTTTC 734  
Db  
966 TAAGCAAGTGTGACAGGAGTGTCTCTGCTCTCAGACTGCGAGGTCCTGATGTTTC 1025  
Qy  
735 AGTTCTCTGACAGCTAACAGTCAATTTGGACTCGAGAGGAATATTTCTAATTTAAAGCT 794  
Db  
1026 AATTCCTGTGCCAGCTAACAGCTCTTTTGGACTCGAGAGGAATATTTCTGACTTAAAGCT 1085  
Qy  
795 CCCAGAGCATCTCATAAAGCTGTGTCATTTGAAGCAATTTATGTGGTGTGTGAGGATA 854  
Db  
1086 TCCAGAGCATCTCATAAAGCTGTGTCATTTGAAGCAATTTATGTGGTGTGTGAGGATA 1145  
Qy  
855 TATGTTCAACCACTCAGATTAATACATGTTTCTAGCGTATGACTTGTCTTAGGAGTG 914  
Db  
1146 TATGTTCAACCACTCAGATTAATACATGTTTCTAGCGTATGACTTGTCTTAGGAGTG 1205  
Qy  
915 GCTTCCACTAAACGTTCTGTGAACAAATGTGGTGTGTGAGATGCTATCTTTGGCAT 974  
Db  
1206 GCTTCCACTAAACGTTCTGTGAACAAATGTGGTGTGTGAGATGCTATCTTTGGCAT 1265  
Qy  
975 ATACAGGATTAATTTTACATGATGAGGAAATTTGATTTCAACTGGGAATGTGACCAA 1034  
Db  
1266 ACATAGGATTAATTTTACATGATGAGGAAATTTGATTTCAACTGGGAATGTGACCAA 1325  
Qy  
1035 TGAGTTGAGAGTATTTTACATTAATGAGTCAATGAGTGTGTGACCCCTTAAGCCAA 1094  
Db  
1326 CGAGCTGAGAGTATTTTACATTAATGAGTCAATGAGTGTGTGACCCCTTAAGCCAA 1385  
Qy  
1095 GGAGCAGTATGCACTGTGGGACTCTGCACATTTGATTTCACTGAGAGTGGCGAGT 1154  
Db  
1386 GGATCAGTATGCACTGTGGGACTCTGCACATTTGATTTCACTGAGAGTGGCGAGT 1445  
Qy  
1155 GGTCACTGTGCTATCTTTGCTCAGTCCCTCTCTATGAGATATATAGCAATGTGAGGA 1214  
Db  
1446 GGTCACTGTGCTATCTTTGCTCAGTCCCTCTCTATGAGATATATAGCAATGTGAGGA 1505  
Qy  
1215 ATATGATTTGATTAAGAACATGAGGATATATTACACACCCAGGTCCTTGTGCAAG 1274  
Db  
1506 ATATGATTTGATTAAGAACATGAGGATATATTACACACCCAGGTCCTTGTGCAAG 1565  
Qy  
1275 GGGTTACGGCCATAGCAGTGTGTTACGACCATAGACCCAGGTCCTTATATGATGTTG 1334  
Db  
1566 GGGTTATGGCCATAGCAGTGTGTTACGACCATAGACCCAGGTCCTTATATGATGTTG 1625  
Qy  
1335 CTACAGGCTTTAGTGCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGT 1394

1626 CTACAGGCTTTAGTGCCAAATAGTACCGGCTTGCAGATGACCTCTACAGATACCATGT 1685  
Qy  
1395 GGATACCCAGATGTGGACCATTTCTTAGGACAGCCGATTTTTCGTTACTTGCACAGC 1454  
Db  
1686 GGACACTCAGATGTGGACCATTTCTTAGGACAGCCGATTTTTCGTTACTTGCATACAGC 1745  
Qy  
1455 TGTGATGTAGTGTGGAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTAT 1514  
Db  
1746 TGTGATGTAGTGTGGAACCATGCTGGTGTGTTGGAGGGAACACACAAATGACACTTCAT 1805  
Qy  
1515 GAGCCATGGCCCAAAATGCTTCTCTCAGATTTTCAATGCTTATGATGCTGTGACCG 1574  
Db  
1806 GAGCCATGGCCCAAAATGCTTCTCTCAGATTTTCAATGCTTATGATGCTGTGACCG 1865  
Qy  
1575 CTGCTCAGTGTCTCCAGACCTGATCTCCACCATGTGTCAACAGATTTGGCCACTCAGC 1634  
Db  
1866 ATGCTCAGTGTCTCCAGACCTGATCTCCACCATGTGTCAACAGATTTGGCCACTCAGC 1925  
Qy  
1635 AGTCTTACACAAACAGCACCATGTATGTGTCGGTGTTCATATAGTCTCTCTCAGCGA 1694  
Db  
1926 AGTCTTACACAAACAGCACCATGTATGTGTCGGTGTTCATATAGTCTCTCTCAGCGA 1985  
Qy  
1695 CATCTGTATTTCACTCGGACAGTGTGATGCGCATCGGAGTGAAGCGCTGTTTATG 1754  
Db  
1986 GGTCTGTATTTCACTCGGACAGTGTGATGCGCATCGGAGTGAAGCGCTGTTTATG 2045  
Qy  
1755 AGCAGGACTGGTATTTGCTGTGTGTGGAACACAGGCTGCTCAGTGTATCTCGTGGC 1814  
Db  
2046 AGCAGGACTGGTATTTGCTGTGTGTGGAACACAGTGTGCTCGATGTACCTCCTGGG 2105  
Qy  
1815 GCTGCAATCTGATGAACAGGAAGAAAGTTAAATCAGATGTTTTCCTTCCAAAGACTCT 1874  
Db  
2106 GTTGCCTCTGATGAACAGGAAGAAAGTTAAATCAGATGTTTTCCTTCCAAAGACTCT 2165  
Qy  
1875 TGACCATCAGAGTGTGACAGCACAGATTTGTACAGCTGTCACAGCAACACCAATGA 1934  
Db  
2166 TGACCATCAGAGTGTGACAGCACAGATTTGTACAGCTGTCACAGCAACACCAATGA 2225  
Qy  
1935 CTGCCACTGGTGCATACCATTTGTGTCGCCAGGAAACACAGCTGCTCAGAGGCCAGAT 1994  
Db  
2226 CTGCCACTGGTGCATACCATTTGTGTCGCCAGGAAACACAGCTGCTCAGAGGCCAGAT 2285  
Qy  
1995 CTCCATTTTATGATGATGAGATTTGCCCAAGGATAACCTATGTACTACTGTAAACAGAA 2054  
Db  
2286 CTCCATTTTATGATGATGAGATTTGCCCAAGGATAACCTATGTACTACTGTAAACAGAA 2345  
Qy  
2055 GACAGCTGCGAGGAGTGTGCCCTGGACCCAGAACTGCCAGTGGAGGCCCGGAATCAGGA 2114  
Db  
2346 AACAGCTGCGAGGAGTGTGCCCTGGACCCAGAACTGCCAGTGGAGGCCCGGAATCAGGA 2405  
Qy  
2115 GTGCATTTGCCCTGCCGGAATATCTGTGGCATTTGCTGGCATTTGTTGGAATCTCATG 2174  
Db  
2406 GTGCATTTGCCCTGCCGGAATATCTGTGGCATTTGCTGGCATTTGTTGGAATCTCATG 2465  
Qy  
2175 TTTGAAAATTAATCTGTCGCAAGGAGATTTATGACAAATGCTAAATTTGTTCTGTAGAAACCA 2234  
Db  
2466 TCTGAAATTAATCTGTCGCAAGGAGATTTATGACAAATGCTAAATTTGTTCTGTAGAAACCA 2525  
Qy  
2235 CAATGCCCTTTTGGTCTCTTTCAACCCAGAGAGGTAGAAATTTGCTTCTTAGAGCAGT 2294  
Db  
2526 CAATGCCCTTTTGGTCTCTTTCAACCCAGAGAGGTAGAAATTTGCTTCTTAGAGCAGT 2585  
Qy  
2295 GCGAATATGACATCTCAGAGCATGTCCAAAGTCACTTAAACCCCATGGTGGCCCT 2354  
Db  
2586 TCGAATATGACATCTCAGAGCATGTCCAAAGTCACTTAAACCCCATGGTGGTCT 2645  
Qy  
2355 TCGGAAGATCAATGTGCTCTAATGCTGCTGGAAGATATGTCCTCCATTTACAAATAGTT 2414  
Db  
2646 TCGGAAGATCAATGTGCTCTAATGCTGCTGGAAGATATGTCCTCCATTTACAAATAGTT 2705  
Qy  
2415 ACTACAGTGGATGCGTCTGAGCCAGTGTGATGCTGGAATTTCTGTGGAATTTTATCAGAAC 2474







treatment of body weight disorders.

Claim 1; Fig 2a; 188pp; English.

This sequence represents a murine mahogany gene of the invention. The mahogany genes are used: (i) to produce recombinant mahogany (mg) proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming therapeutics; (iii) as a source of diagnostic probes and primers for detecting expression of mg genes or mutations, regulatory defects, in this gene, or for isolation of related sequences; and (iv) in (cell-based) gene therapy. (II) are used to raise specific antibodies (Ab); to identify other (extra)cellular products involved in weight regulation, and to screen for agents that disrupt interaction between (II) and other macromolecules. The Ab are used to detect abnormal levels (or function) of (II) for diagnosis, prognosis or monitoring of treatment); to evaluate (II)-expressing cells intended for cell therapy, and as therapeutic mg inhibitors. Cells that express the mg gene (or contain the mg polypeptide) are used to identify agents (A) that modulate mg activity. (A) are potentially useful for the treatment of body weight disorders, particularly obesity, cachexia or anorexia, or other conditions associated with the mg gene such as hyperpigmentation, hyperphagia and disorders that result in increased metabolic rate

Sequence 8827 BP: 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;

Query Match 81.7%; Score 3324.8; DB 3; Length 8827;  
Best Local Similarity 91.1%; Pred. No. 0;  
Matches 3533; Conservative 0; Mismatches 347; Indels 0;

189	QY	ACTAACTGGATCTCTCGGTTTGTGCAGATGGACCTGGAAATATATAAATACAAAACGAA	248
677	Db	ACTAACTGGCTCTCTCGAATTGTACAGATGGACCTGGGAAATTATAAATATAGACGAA	736
249	QY	GTGCACGTGGCTCAATTGAAGACGACGCAAAATAGAAATAATGAGACTTCGTTTCAATCAATT	308
737	Db	GTGCACATGGCTCAATTGAAGACGACGCAAAATAGAAATAATGAGACTTCGCTTCAACCAATT	796
309	QY	TGCTACAGAGGTAGATTGGGACCAATTATATGTTTATGATGGGAGACTCAATTTATGCACC	368
797	Db	TGCTACAGAAATGACTGGGACCAATTATATGTTTATGATGGGAGACTCAATCTACGCACC	856
369	QY	GCTAGTTGCTGCATTTATAGTGCCCTCAATTGTTCTGTAGAGAGATGGCAATGAGACTGTCCC	428
857	Db	TCTGATTGCTGCCCTTTATGTCGCCCTCAATTGTTCTGTAAAGAGATGGCAATGAGACGGCTCC	916
429	QY	TGAGTTGTTTGCCACATCAGCTTATGCTTGTCTGCATTTTATTAGTCATGCTGCTTATAA	488
917	Db	TGAGGTCACTGTCACTTCAGGTTATGACTGCTGCAATTTTTCAGTGAATGCTGCTTATAA	976
489	QY	TTTGACTGGAAATTAATAATTACTTACAGTTTGTATGTGTGCCAAATAAATGCTCAGGCGG	548
977	Db	TCTGTACTGGAAATTAATATCATCTTACAATTTGACATGTGCCGAATAAATGCTCAGGCGG	1036
549	QY	AGGAGAGTGTAAGATCAGTAAATAGCAGCGATCTGTTGAATGTGAATGTTCTGAAAACTG	608
1037	Db	AGGAGAGTGTAAGACGAGTAACACGACGCGCTGTTGAGTGTGAATGTTCTGAAAACTG	1096
609	QY	GAAGGTGAAGCATGTGCATTCCTCACTGTACAGACAACTGTGTGTTTCTCTCATCGAGG	668
1097	Db	GAAGGGGAGTGTGATCTCTCACTGTACAGAACTGTGGCTTTCTCTACCGAGG	1156
669	QY	CATCTGCAATTCAGTGTGATGTACAGGATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGG	728
1157	Db	CATCTGTAATGCAAGCGATACGAGGGTGTCTCTGCTTCTCACTGGCAGGGTCTCTGG	1216
729	QY	ATGTTTCAGTTCCTGTACCGCTACACAGTCAATTTTGGACTCGAGAGAAATTTCTAACTT	788
1217	Db	ATGTTTCAATTCCTGTGCAGCTACCAAGTCTTTTGGACTCGAGAGAAATTTCTGATTT	1276
789	QY	AAAGCTCCCCAGACATCTCATAAAGCTGTGGTCAATGGGAAACAATTAATGTGGGTGTGG	848
1277	Db	AAAGCTTCCAGAGCCCTCTCATAAAGCTGTGGTCAATGGGAAATAAATGTGGGTGTGG	1336

2417	DB	CAATGACTGCCATGGTGCAATGATCACTGTGTCCTGTGAACCAACAGCTGCACGAAGG	2476
1989	QY	CCAGATCTCCATTTTATAGGTATGAGAAATGCCCCAAAGATAAACCCTATGTACTACTGTAA	2048
2477	DB	CCAGATCTCCATTTGCCAAGTATGAGAGTTGCCCCCAAGGATAACCCCATGTACTACTGC	2536
2049	QY	CAAGAGACAGAGCTGCAGGAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGAGAGCCCGGAA	2108
2537	DB	TAAAGAAACAGAGCTGCAGGAGAGCTGTGCCCTTAGACCAAGAACTGCCAGTGGAGAGCCCGGAA	2596
2109	QY	TCAGGAGTGCATTTGCCCTGCCCGAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAA	2168
2597	DB	TCAGAGTGCATTTGCCCTGCCCGAAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGGAAA	2656
2169	QY	CTCATGTTTGAAATTAATACTACTGCCAAGGAGAAATATGACAATGCTAAATTTGTTCTGTAG	2228
2657	DB	CTCGTGTCTGAAAATCACTACTGCTAAGGAGAAATATGACAATGCTAAATTTGTTCTGTAG	2716
2229	QY	GAACCAAAATGCCCTTTTGGCTTCTCTTCAAAACCCAGAGAAGGTAGAAATTTCTCCCTTAA	2288
2717	DB	GAACCAAAATGCCCTTTTGGCTTCTCTTCAAAACCCAGAGAAGGTAGAAATTTCTCCCTTAA	2776
2289	QY	GCAGCTGCGAATAATGCAGTCACTCTCAGAGCATGTCCAAGCTCACTCTAACCCCATGGGT	2348
2777	DB	GCAGCTTCGATTAATGCAATCACTCAAGATATGTCTCCAAGCTCACTCTAACCCCATGGGT	2836
2349	QY	CGGCCCTCGGAAGATCAATGTGCTTACTGGTCTGGAGAGATATGTCCCATTTACAAA	2408
2837	DB	TGCTGTTTCGGAAGATCAATGTGCTTACTGGTCTGGGAGGATATGTCCATTTACAAA	2896
2409	QY	TAGTTTACTACAGTGGATGCCGCTGTGAGCCAGTGATCTGCGAATTTCTGTGGAAATTTTATC	2468
2897	DB	TAGTTTGTCTCAGTGGATGCCATCTGAGCCAGTGATGCTGGCTTCTGTGGGATCTGTCTC	2956
2469	QY	AGAAACCAAGTACTCGGGGACTGAAGGCTGCAAACCTTGCAATCAACCCACTCAATGGTAGTGT	2528
2957	DB	AGAGCCTTAGTACTCGGGGATTAAGGCTGCAAACCTTGCAATCAACCCCTCAATGGCAGCGT	3016
2529	QY	CTGTGAAGGCCCTGCAAAACCAACAGTGTCTAAGCAGTGGCCGACACCAATGTGCCCTTGAGGAC	2588
3017	DB	CTGTGAAGGCCCTGCAAAACCAACAGTGTCCAAGCAGTGGCCGACACCAATGTGCCCTTGAGGAC	3076
2589	QY	AGCATGTGGAGATTGCACAGCGCAGCTCTGATGTGCATGTGGTGCAGCAACATGAAGCA	2648
3077	DB	AGCGTGTGGCGAGTGCACTAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATGAAGCA	3136
2649	QY	GTGTGTGGAATCCAAATGCCCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTA	2708
3137	DB	GTGTGTGGAATCCAAATGCCCTACGTGGCTCTCTCCCTTTTGGCCAGTGTATGGATGGTA	3196
2709	QY	TACGATGAGCACCTTGCCGCCCTGAAAATTTGTTCAAGCTACTGTACCTGTAGTCAATTTGCTT	2768
3197	DB	TACGATGAGCAGCTGCCCACTGAAATTTGCTCTGGCTACTGTACCTGCAGCCATTTGCTT	3256
2769	QY	GGAGCAACAGGCTGTGGCTGGTGTAATGATCCAGCAATACTTGGCAAGGGAATATGCAT	2828
3257	DB	GGAGCAGCAGGCTGTGGTGGTGTAATGATCCAGCAATACTTGGCAAGGGAATATGAT	3316
2829	QY	AGAGGGTTTCTATAAAGGACCAAGTGAAGATGCCCTTGCACAGCCCTTACAGAAATTTCTA	2888
3317	DB	TGAGGGCAGCTATAAAGGACCTGTGAAGATGCCCTTACAGCCCTCTCCAGGAATGTGTA	3376
2889	QY	TCCACAGCCCTGTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTT	2948
3377	DB	TCCACAGCCCTTCTGAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTT	3436
2949	QY	CATTCACTGTCCAGCTTGGCAATGCAAGGCCACAGTAATGCATCAATCAGAGCATCTG	3008
3437	DB	CATTCACTGTCCAGCTTGGCAATGCAAGGCCACAGTAATGCATCAATCAGAGCATCTG	3496
3009	QY	TGAGAAGTGTGAGAACTTGACCAAGGCAAGCACTGCGAGACCTTGCATATCTGGCTTCTA	3068

RESULT 12  
AAS72660

ID AAS72660 standard; cDNA; 3490 BP.  
XX AAS72660;  
AC  
XX  
XX 13-FEB-2002 (first entry)  
DT  
XX  
XX DNA encoding novel human diagnostic protein #8464.  
DE  
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
KW  
XX Homo sapiens.  
OS  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX  
XX P-PSDB; ABG08473.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
XX Claim 1; SEQ ID NO 8464; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;  
SQ  
Query Match 69.1%; Score 2810.4; DB 5; Length 3490;  
Best Local Similarity 94.3%; Pred. No. 0;  
Matches 3236; Conservative 0; Mismatches 51; Indels 145; Gaps 25;  
187 AGACTAACTGGATCTTCTGGTTGTGACAGATGGACCTGGAAATTAATAACAAACG 246  
22 AGACTAACTGGATCTTCTGGTTGTGACAGATGGACCTGGAAATTAATAACAAACG 81  
247 AAGTGACGCTGGCTCAATTAAGGACGACCAATAGAAATATGAGACTTCGTTTCAATCAT 306  
82 AAGTGACGCTGGCTCAATTAAGGACGACCAATAGAAATATGAGACTTCGTTTCAATCAT 141  
307 TTGCTACAGAGTGTAGTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCA 366

142 TTGCTACAGAGTGTAGTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCA 201  
357 CCGTAGTGTGCTGCAATTTAGTGGCTCATTTGTTCTCTGAGAGAGATGCAATGAGACTGTC 426  
202 CCGTAGTGTGCTGCAATTTAGTGGCTCATTTGTTCTCTGAGAGAGATGCAATGAGACTGTC 261  
427 CTTGAGGTTGTTGCCACATCAGGTTATGCTTATGCTGCAATTTTTTCTGATGATGCTGCTTAT 486  
262 CTTGAGGTTGTTGCCACATCAGGTTATGCTTATGCTGCAATTTTTTCTGATGATGCTGCTTAT 321  
487 AATTGACTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 546  
322 AATTGACTGCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 381  
547 CGAGAGAGTGTAAAGATCAGTAATAGCAGGATAGTGTGAATGTAATGTTCTGAAAC 606  
382 CGAGAGAGTGTAAAGATCAGTAATAGCAGGATAGTGTGAATGTAATGTTCTGAAAC 441  
607 TCGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGCAACATGTTGTTTCTCTCATCGA 666  
442 TCGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGCAACATGTTGTTTCTCTCATCGA 501  
667 GGCATCTGCAATTCAGTGTATGCTGAGGATGCTCTGCTTCTCAGACTGCGAGGCTCT 726  
502 GGCATCTGCAATTCAGTGTATGCTGAGGATGCTCTGCTTCTCAGACTGCGAGGCTCT 561  
727 GGATGTTTCAAGTGTATGCTGAGGATGCTCTGCTTCTCAGACTGCGAGGCTCT 785  
562 GGATGTTTCAAGTGTATGCTGAGGATGCTCTGCTTCTCAGACTGCGAGGATGCTCT 621  
786 CTT--AAAGCTTCCCGAGGATGCTCTAATAAGCTGTGTCATGGAACATATATGTTGGTT 843  
622 CTTTAAAGCTTCCCGAGGATGCTCTAATAAGCTGTGTCATGGAACATATATGTTGGTT 681  
844 GTTGAGGATATATGTTTCAACACTCAGATTAATAACATGTTTCTAGCGTATGACCTTGCT 903  
682 GTTGAGGATATATGTTTCAACACTCAGATTAATAACATGTTTCTAGCGTATGACCTTGCT 741  
904 TCTAGGAGTGGCTTCCCACTAAACCGTTCTGTGAAACAAATGTTGTTTGTAGATGTTGCT 963  
742 TCTAGGAGTGGCTTCCCACTAAACCGTTCTGTGAAACAAATGTTGTTTGTAGATGTTGCT 801  
964 TCTTGGCATTAACAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1023  
802 TCTTGGCATTAACAGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 861  
1024 AATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGGTTGTTGAC 1083  
862 AATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGGTTGTTGAC 921  
1084 CTTAAGGCAAGGAGGATGAGTGGTGGGCTGTCACATTCGACACATGTTTACATCAAG 1143  
922 CTTAAGGCAAGGAGGATGAGTGGTGGGCTGTCACATTCGACACATGTTTACATCAAG 981  
1144 AATGGCCAGTGGTTCATGCTGGTTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGC 1203  
982 AATGGCCAGTGGTTCATGCTGGTTCATCTTTGGTCACTGCCCTCTCTATGATATATAAGC 1041  
1204 AATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGGTTGTTGAC 1263  
1042 AATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGGTTGTTGAC 1101  
1264 CTTGTGCAAGGCGGTTACGGCCATAGCAGTGTGTTTACGACCATAGGACCGGCGCTATAC 1323  
1102 CTTGTGCAAGGCGGTTACGGCCATAGCAGTGTGTTTACGACCATAGGACCGGCGCTATAC 1161  
1324 GTTCATGTTGCTCAAGGCTTTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1383  
1162 GTTCATGTTGCTCAAGGCTTTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1221  
1384 CGATATGATGTTGATACCCAGATGTTGACCAATTCCTTAAGGACAGCGGATTTTCCGTTAC 1443



QY	3570	CAAAATTCAGAT	3581	75	GAGCGCGCGCGCACACTGTGTCAACGCGCGTGTGTCAACCCCTGCGCACCCGCGCAGTGGT	134
DB	3334	CAAAATTCAGGT	3345	294	GGATGTGACCGCGCGTGTGTCAACGCGCGTGTGTCAACCCCTGCGCACCCGCGCAGTGGT	353
RESULT 13						
ID	AAZ91922					
XX	AAZ91922	standard; cDNA; 2625 BP.		135	CTGCCCGCGCGTGGGTGGCGGAGCAATGCCAGACACTGGCGGCGCGCTTCAGACTAAC	194
AC	AAZ91922;			354	CTGCCCGCGCGTGGGTGGCGGAGCAATGCCAGACACTGGCGGCGCGCTTCAGACTAAC	413
DT	08-JUN-2000	(first entry)		195	TGGATCTTCTGGGTGGTGTGACAGATGGACCTGGAAATATATAAATACAAAACGAAGTGCAC	254
DE	Human mahogany protein coding sequence #4.			414	TGGATCTTCTGGGTGGTGTGACAGATGGACCTGGAAATATATAAATACAAAACGAAGTGCAC	473
XX	Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;			255	GTGGCTCAATGAAGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTGTAC	314
KW	weight regulation; cell therapy; body weight disorder; cachexia;			315	AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGT	374
KW	anorexia; hyperpigmentation; increased metabolic rate disorder;			534	AGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGT	593
KW	hyperphagia; Antiobesity; antianorexic; anticachexic; ss.			375	TGCTGCATTTAGTGGCCTCATTTGTTCTGAGAGATGGCAATGAGACTGTCCTCGAGGT	434
OS	Homo sapiens.			594	TGCTGCATTTAGTGGCCTCATTTGTTCTGAGAGATGGCAATGAGACTGTCCTCGAGGT	653
XX	WO200005373-A2.			435	TGTTGCCACATCAGTTATGCTTGTGCTGCTATTTTATGATGCTGCTTATATTTGAC	494
PN	03-FEB-2000.			654	TGTTGCCACATCAGTTATGCTTGTGCTGCTATTTTATGATGCTGCTTATATTTGAC	713
XX	21-JUL-1999; 99WO-US016484.			495	TGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	554
PR	21-JUL-1998; 98US-0093630P.			714	TGGATTTAATATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	773
PR	20-OCT-1998; 98US-0104978P.			555	GTGTAAGATCAAGTATAGCAGGATCTCTGTTGAATGTAATGTTCTGAAACTGGAAGG	614
PR	05-FEB-1999; 99US-00245041.			774	GTGTAAGATCAAGTATAGCAGGATCTCTGTTGAATGTAATGTTCTGAAACTGGAAGG	833
XX	(MILL-) MILLENIUM PHARM INC.			615	TGAAGCATGTGACATCTCTCTACTGTACAGACAACTGTTGTTTCTCATCGAGGCATCTG	674
PI	Moore K, Nagle DL;			834	TGAAGCATGTGACATCTCTCTACTGTACAGACAACTGTTGTTTCTCATCGAGGCATCTG	893
XX	WPI; 2000-195103/17.			675	CAATTCAGTGTGATGTCAGAGGATGCTCTGCTTCTCAGACATGGCAGGCTCTGATGTTT	734
DR	P-PSDB; AAY81509.			894	CAATTCAGTGTGATGTCAGAGGATGCTCTGCTTCTCAGACATGGCAGGCTCTGATGTTT	953
XX	New human and murine mahogany genes, useful, e.g. for diagnosis and			735	AGTTCTCTGTACAGTAAACCACTCATTTTGGACTCGAGAGAAATATTTCTAACTTAAAGCT	794
PT	treatment of body weight disorders.			954	AGTTCTCTGTACAGTAAACCACTCATTTTGGACTCGAGAGAAATATTTCTAACTTAAAGCT	1013
PS	Claim 1; Fig 20a; 188pp; English.			795	CCCCAGACATCTCATAAAGCTGTGGTCAATGGAACAACTATGTTGGGTGTTTGGAGGATA	854
XX	This sequence represents a human mahogany gene of the invention. The			1014	CCCCAGACATCTCATAAAGCTGTGGTCAATGGAACAACTATGTTGGGTGTTTGGAGGATA	1073
CC	mahogany genes are used: (i) to produce recombinant mahogany (mg)			855	TATGTTCAACCACTCAGATTAATACATGTTTACGCTTACGCTTACGCTTACGCTTACGCT	914
CC	proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming			1074	TATGTTCAACCACTCAGATTAATACATGTTTACGCTTACGCTTACGCTTACGCTTACGCT	1133
CC	therapeutics; (iii) as a source of diagnostic probes and primers for			915	GCTTCCACTTAAACCGTCTCTGTAACAACTGTTGTTAGATATGTTGTTGTTGTTGTTGTTG	974
CC	detecting expression of mg genes or mutations, regulatory defects, in			1134	GCTTCCACTTAAACCGTCTCTGTAACAACTGTTGTTAGATATGTTGTTGTTGTTGTTGTTG	1193
CC	this gene, or for isolation of related sequences; and (iv) in (cell-			975	ATACAGGATATAATTTACATGTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1034
CC	based) gene therapy. (II) are used to raise specific antibodies (Ab); to			1194	ATACAGGATATAATTTACATGTTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1253
CC	identify other (extra)cellular products involved in weight regulation,			1035	TGAGTTGAGAGTATTTTACATGTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT	1094
CC	and to screen for agents that disrupt interaction between (II) and other			1254	TGAGTTGAGAGTATTTTACATGTTATGAGTATGAGTATGAGTATGAGTATGAGTATGAGTAT	1313
CC	macromolecules. The Ab are used to detect abnormal levels (or function)			1095	GGAGCAGTATGAGTGTGGGACCTCTGTCACATTTGTTTACACTGAAGATGCGGAGT	1154
CC	of (II) (for diagnosis, prognosis or monitoring of treatment); to			1314	GGAGCAGTATGAGTGTGGGACCTCTGTCACATTTGTTTACACTGAAGATGCGGAGT	1373
CC	evaluate (II)-expressing cells intended for cell therapy, and as			1155	GGTCAATGCTGCTCATCTTTGGTCACTGCTCTCTGATGATATATAAGCAATGTCAGGA	1214
CC	therapeutic mg inhibitors. Cells that express the mg gene (or contain the					
CC	mg polypeptide) are used to identify agents (A) that modulate mg					
CC	activity. (A) are potentially useful for the treatment of body weight					
CC	disorders, particularly obesity, cachexia or anorexia, or other					
CC	conditions associated with the mg gene such as hyperpigmentation,					
CC	hyperphagia and disorders that result in increased metabolic rate					
XX	Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;					
SQ	Query Match 50.4%; Score 2051.8; DB 3; Length 2625;					
	Best Local Similarity 97.1%; Pred. No. 0;					
	Matches 2089; Conservative 0; Mismatches 62; Indels 0; Gaps 0;					
QY	15	GGCGGCAACTGAGCAAGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	74			
DB	234	GGAGCGGAGCCCGCTCGCGGTGGCGGCGCGGTGTCGCGGTGCGCGCGAGCGGAGGAGGAG	293			

```
Db 1374 GGTGATGCTGGTCACTTTGGTCACTGGCCCTCTCTATGATATATAAGCAATGTCAGGA 1433
QY 1215 ATATGATTTGGATAAGAACACATGAGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1274
Db 1434 ATATGATTTGGATAAGAACACATGAGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGG 1493
QY 1275 GGGTTACGCCCATAGCAGTGTGTTTACGACCATAGGACCGGCCCTATAGTTCATGGTGG 1334
Db 1494 GGGTTACGCCCATAGCAGTGTGTTTACGACCATAGGACCGGCCCTATAGTTCATGGTGG 1553
QY 1335 CTACAGAGGCTTTTCAGTGCCTCAATAGTACCGGCTTGCGAGATGATCTCTACCGATATGATGT 1394
Db 1554 CTACAGAGGCTTTTCAGTGCCTCAATAGTACCGGCTTGCGAGATGATCTCTACCGATATGATGT 1613
QY 1395 GGNATCCAGATGTCACCATCTTTAAGGACACCGGATTTTCCGTTACTTGCACACAGC 1454
Db 1614 GGNATCCAGATGTCACCATCTTTAAGGACACCGGATTTTCCGTTACTTGCACACAGC 1673
QY 1455 TGTGATAGTGAGTGGAAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTAT 1514
Db 1674 TGTGATAGTGAGTGGAAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTAT 1733
QY 1515 GAGCCATGGCGCAAAATGCTTCTCTTCAGATTTTCATGGCTATGATGATTCGCTGTGACCG 1574
Db 1734 GAGCCATGGCGCAAAATGCTTCTCTTCAGATTTTCATGGCTATGATGATTCGCTGTGACCG 1793
QY 1575 CTGGTCAGTGCTTCCAGACCTGATCTCCACATGATGTCACAGATTTGGCCATTCAGC 1634
Db 1794 CTGGTCAGTGCTTCCAGACCTGATCTCCACATGATGTCACAGATTTGGCCATTCAGC 1853
QY 1635 AGTCTTTACCAACAGCACCACCATGATGTTGTTGGTGGTTCATAGTCTCTCTCTCAGCGA 1694
Db 1854 AGTCTTTACCAACAGCACCACCATGATGTTGTTGGTGGTTCATAGTCTCTCTCTCAGCGA 1913
QY 1695 CATCTCTGTTATTCACCTCGGAACAGTGTGATGCGGATCGGAGTGAAGCGCTTGTAGC 1754
Db 1914 CATCTCTGTTATTCACCTCGGAACAGTGTGATGCGGATCGGAGTGAAGCGCTTGTAGC 1973
QY 1755 AGCAGGACCTGGTATTTCCGTTGTGTGGAACACAGGCTGCTCTCAGTGTATCTCGTGGGC 1814
Db 1974 AGCAGGACCTGGTATTTCCGTTGTGTGGAACACAGGCTGCTCTCAGTGTATCTCGTGGGC 2033
QY 1815 GGTGGCAACTGATGAAACAGAAAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCT 1874
Db 2034 GGTGGCAACTGATGAAACAGAAAGAAAGTTAAATCAGAAATGTTTTCAAAAGAACTCT 2093
QY 1875 TGACCATGACAGATGTGACAGCACACAGATTTGTACAGTGCACAGCCAAACCAATGA 1934
Db 2094 TGACCATGACAGATGTGACAGCACACAGATTTGTACAGTGCACAGCCAAACCAATGA 2153
QY 1935 CTGCCACTGGTCAATGACCAATGTTGTCGCCAGGAACCAACAGTGTCTAGAGGGCCAGAT 1994
Db 2154 CTGCCACTGGTCAATGACCAATGTTGTCGCCAGGAACCAACAGTGTCTAGAGGGCCAGAT 2213
QY 1995 CTCATTTTATGATGAGATTTGCCCAAGGATACCTCTACTACTGTAAACAGAA 2054
Db 2214 CTCATTTTATGATGAGATTTGCCCAAGGATACCTCTACTACTGTAAACAGAA 2273
QY 2055 GACCACTGTCAGGAGTGTGCCCTGGACACAGAACTGCCAGTGGGAGCCCGGAATCAGGA 2114
Db 2274 GACCACTGTCAGGAGTGTGCCCTGGACACAGAACTGCCAGTGGGAGCCCGGAATCAGGA 2333
QY 2115 GTGCATTTGCCCTGCGCAAAATATCTGTGGCATTGGCTGGCATTTGGTGG 2165
Db 2334 GTGCATTTGCCCTGCGCGGTTCGAGGGTCACTTTGGTGTGTGGG 2384
```

## RESULT 14

AAZ91919

ID AAZ91919 standard; cDNA; 6373 BP.

XX

AC AAZ91919;

```
XX 08-JUN-2000 (first entry)
DE Human mahogany protein coding sequence #1.
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
OS Homo sapiens.
XX Key Location/Qualifiers
FH misc_difference 448
FT /*tag= a
FT /note= "represented in specification as : "
FT misc_difference 615
FT /*tag= b
FT /note= "represented in specification as : "
FT misc_difference 683
FT /*tag= c
FT /note= "represented in specification as : "
XX WO200005373-A2.
XX 03-FEB-2000.
XX 21-JUL-1999; 99WO-US016484.
XX 21-JUL-1998; 98US-0093630P.
XX 20-OCT-1998; 98US-0104978P.
XX 05-FEB-1999; 99US-00245041.
XX (MILL-) MILLENIUM PHARM INC.
XX Moore K, Nagle DL;
XX WPI; 2000-195103/17.
XX P-PSDB; AAY81906.
XX New human and murine mahogany genes, useful, e.g. for diagnosis and
XX treatment of body weight disorders.
XX Claim 1; Fig 10a; 188pp; English.
XX This sequence represents a human mahogany gene of the invention. The
XX mahogany genes are used: (i) to produce recombinant mahogany (mg)
XX proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
XX therapeutics; (iii) as a source of diagnostic probes and primers for
XX detecting expression of mg genes or mutations, regulatory defects in
XX this gene, or for isolation of related sequences; and (iv) in (cell-
XX based) gene therapy. (II) are used to raise specific antibodies (Ab); to
XX identify other (extra)cellular products involved in weight regulation,
XX and to screen for agents that disrupt interaction between (II) and other
XX macromolecules. The Ab are used to detect abnormal levels (or function)
XX of (II) (for diagnosis, prognosis or monitoring of treatment); to
XX evaluate (II)-expressing cells intended for cell therapy, and as
XX therapeutic mg inhibitors. Cells that express the mg gene (or contain the
XX mg polypeptide) are used to identify agents (A) that modulate mg
XX activity. (A) are potentially useful for the treatment of body weight
XX disorders, particularly obesity, cachexia or anorexia, or other
XX conditions associated with the mg gene such as hyperpigmentation,
XX hyperphagia and disorders that result in increased metabolic rate
XX
SQ Sequence 6373 BP; 1736 A; 1468 C; 1460 G; 1699 T; 0 U; 10 Other;
Query Match 45.0%; Score 1830.2; DB 3; Length 6373;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 1986; Conservative 7; Mismatches 41; Indels 26; Gaps 11;
QY 2035 ATGTACTACTGTATACAGAACAGACAGCTGCAGGAGCTGTGCCCTGGACCAAGTCCAG 2094
Db 1 ATGTACTACTGTATACAGAACAGACAGCTGCAGGAGCTGTGCCCTGGACCAAGTCCAG 60
```







DE Murine homologue of attractin/mahogany (HAM) polypeptide encoding cDNA.  
 KW HAM; homologue of attractin/mahogany; immunosuppressive; cytostatic;  
 KW antiinflammatory; cardiant; osteopathic; gene therapy; mouse; gene; ss.  
 XX Mus sp.  
 XX Key Location/Qualifiers  
 XX CDS 1..4134  
 XX /tag= a  
 XX /product= "HAM"  
 XX /note= "homologue of attractin/mahogany"  
 XX WC020297120-A1.  
 XX PD 05-DEC-2002.  
 XX PF 23-MAY-2002; 2002WO-US016391.  
 XX PR 25-MAY-2001; 2001US-0293608P.  
 XX PR 24-SEP-2001; 2001US-0324626P.  
 XX PA (IMV ) IMMUNEX CORP.  
 XX PI Anderson DM;  
 XX DR WPI; 2003-140486/13.  
 XX DR P-PSDB; ABP71102.  
 XX PT New Homologue of Attractin/Mahogany (HAM) polypeptide, useful for  
 XX PT treating HAM-associated disorder consisting of inflammatory, autoimmune,  
 XX PT cell proliferative or cardiovascular disorders.  
 XX PS Claim 14; Fig 5A-D; 89pp; English.  
 XX CC The invention relates to Homologue of Attractin/Mahogany (HAM)  
 CC polypeptides and encoding polynucleotides. The HAM polypeptides can be  
 CC expressed by standard recombinant methodology. The HAM polypeptides are  
 CC useful for treating HAM-associated disorder consisting of inflammatory,  
 CC autoimmune, graft-versus-host, neurological, myelination, cell  
 CC proliferative, cardiovascular, haematologic, liver, metabolic, weight or  
 CC bone disorder. The present sequence represents a cDNA encoding a mouse  
 CC HAM polypeptide  
 XX SQ Sequence 4134 BP; 1085 A; 981 C; 1036 G; 1032 T; 0 U; 0 Other;  
 Query Match 36.4%; Score 1481; DB 7; Length 4134;  
 Best Local Similarity 63.4%; Pred. No. 0;  
 Matches 2507; Conservative 0; Mismatches 1360; Indels 90; Gaps 12;  
 QY 117 TGGCAGCGGCGTCTGCTGCGCGCGGCTGGGCGGCGATGCCAGCTGGG 176  
 DB 219 TGTCAACTCCACTGCTGTGGACCGGCGCTGGGCGGCGATGCCAGCTGGC 278  
 QY 177 GGGCGGCTTCAGACTAATCGATCTCTGGGTTGTGACAGATGGACCTGGAAATATATA 236  
 DB 279 GGGCAGGTTCAAGTAAACAGACCTCTGATATTAACAGATGGACCAATTAATATA 338  
 QY 237 ATCAAAACGAAGTCAGCTGGCTCATTTGAAGACAGCAATAGAAATAGAGCTTCG 296  
 DB 339 ATATAAAACAAAGTACATGGTGTAAATGAAGGCTATCCAAATGCAAGTGAAGTTAAG 398  
 QY 297 TTTCAATCATTTTGTACAGAGTGTAGTTGGGACCATTTATATGTTTATGATGGGACATC 356  
 DB 399 ATTCAATCATTTTGTACAGATGGAGCTGGGATCATATGATGTTTATGATGGAGTTTC 458  
 QY 357 AATTATGACCGCTAGTTGCTGCAATTTAGTGGCTCATTTGTTCTCGAGAGATGGCAA 416  
 DB 459 TATATACGACCTTTAGTAGCTGCTACTTATAGTGGCTTGTATGCTTCTCTGAAGTGGGGTAA 518  
 QY 417 TGAGCTGCTCCCTGAGTTGTTGCCATCATCAGCTATGCTTGTCTCATTTTATAGTGA 476  
 DB 519 CGAGACCGTCCCTGAGGTGTCAGAGCTGTGGCTACGGCTCTCCATCTTTTACGCGA 578

QY 477 TGGCTCTATAAATTTGATGTGATTTAATATTACTTACATTTTATGATATGTGCTCAATAA 536  
 DB 579 TGGTGCATATAACCTAACTGGCTTCAACATTTTATTCGATCAATTCCTGCTCAACA 638  
 QY 537 CTGCTCAGGCGGAGAGAGTGTAAAGATCAGTAATAGCAGCGATACT---GTTGAATGTGA 593  
 DB 639 CTGCTCTGGTCATGGAAGTGTACAACACAGTGTCTCTGTTGCAAGTCAAGTGTATTGTGA 698  
 QY 594 ATGTTCTCAAAACCTGGAAGGTGAAGCATGTGACATTCCTCAGTGTACAGCAACTGTGG 653  
 DB 699 ATGGCAAAATATCTGGAAGGAGCATGTGACATTCCTTACTGTAAAGCAATTTGGTGG 758  
 QY 654 TTTTCTCATCGAGGCTCTGCAATTCAGTGTGAGAGGATGCTCTGCTCTCTCAGA 713  
 DB 759 GAGTCCAGATCATGGCTACTGTGACCTAACAGGAGAGAAACTCTGTCTGCAACGATAG 818  
 QY 714 CTGGCAGGCTCTGGATGTTTCTGATCTGACCTAACAGCTAACAGCTCAATTTGGACTCGA 773  
 DB 819 TTGGCAAGGCCAGATGTTTCTCTGAATGTCCTTCTACTGAGTCTTACTGGATTTTGCC 878  
 QY 774 GGAATATTCTAACTTAAAGCTCCCTCCGTTAGTTCGGGCTCACATAAGCAGTTTACATGG 827  
 DB 879 AAATGTTAAACCTTTCAGCCCTTCGGTAGTTCGGGCTCACATAAGCAGTTTACATGG 938  
 QY 828 AAACATATGTTGGTGTGAGAGATATATGTTCAACGACCTCAGATTAACATGTTCT 887  
 DB 939 GAAATTCATGTTGGTGTGAGAGATATATCTTTAACTACAGTTCCTTTTCAATGTTCT 998  
 QY 888 AGCGTATGACCTTCTCTAGGAGTGGCTTCCACTAAACCGTCTCTGTAACAAATGTGGT 947  
 DB 999 GAATTAACATTTAAGAACAGTATATGGAATGATGAGTGTCTGATCAAGGGGCCCT---CT 1055  
 QY 948 TGTAGATATGTTCAATCTTTGGCATTAACAAGATAAATAATTTACATGATGAGAGAA 1007  
 DB 1056 TCAGAGATACGACATTTCTTGGCCCTTATACAGAAACATCTTTTATGACGAGGCGAG 1115  
 QY 1008 AATTGATTCAA---CTGGATGTGACCAATGAGTGTGAGAGTGTTCACATTCATAATGA 1064  
 DB 1116 AATGAGACCAAGTATGCAACGCTCCGATGAATATGAGTGTGTTTACGTACCGACTCA 1175  
 QY 1065 GTCATGTTGTTGTTGACCCCTTAAGGCAAGG-----AGCAGTATGAGTGGTGG 1115  
 DB 1176 ATCATGAGACGAGAAACCCCTCTGCTTGGCCACAGTCAGCAGTACGCTGTGGAGGG 1235  
 QY 1116 GCATCTGACACATTTGTATACATGAAGATGCGCGAGTGTGATGCTGTCATCTTTGG 1175  
 DB 1236 ACCTCGGCACACATCATGAGCTGGAGCTGGACAGTAGACGCTGTCATGATTTGCTATT 1295  
 QY 1176 TCATGCTCCTCTCTATGATATATAAGCAATGTGCAAGATATGATTTGGATAAGAACAC 1235  
 DB 1296 ATATTTCCATATATGCTATACAGCAGCATACAGNATACCATATCTCTCTCAACAC 1355  
 QY 1236 ATGAGATATATACACCCAGGTTGCCCTTGTGCAAGGGGTTTACGGCCATACAGTGT 1295  
 DB 1356 TTGGTGTAGTTCCAGAAACGAAAGGAGCCATTTGTCAAGGTGGAATATGCGCCACAAAGT 1415  
 QY 1296 TTACAGCATAGACAGGCGCCCTATAGTTCATGTTGGCTTACAGGCTTTTCAAGTGC 1355  
 DB 1416 GTATGACGAATGACCAAGTCCATCTATGTTTCCAGAGGCTTACAAAGATTTGCTTGGCA 1475  
 QY 1356 TAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACAT 1415  
 DB 1476 TAAGTACCGGCTTGTGGACGACCTCTATAAGTAAAGTCAACACAGGACTTGGACTAT 1535  
 QY 1416 TCTTAAGACAGCCGATTTTTCGGTTACTTGCACACAGCTGTGATGAGTGGAGCAAT 1475  
 DB 1536 TTTGAAGAAAGTGGGTTTGGCAGATACCTTTCACTCAGCTCTTCTTATCAATGAGCTAT 1595  
 QY 1476 GCTGTGTTTGGGGAACACACAAATGACATCTATGAGCCATGCGGCCAATGCTT 1535  
 DB 1596 GCTTATTTTGGAGGAATACCCATATGATATCTTCCCTGAGTAAACGTTGCAAAATGTTT 1655

1536	Qy	CTCTTCAGATTTCATGCGCTATGACATTGCTGTGACCGCTGTCAGTCTGCTTCCAGACC	1599
1537	Qy		
1538	Qy		
1539	Qy		
1540	Qy		
1541	Qy		
1542	Qy		
1543	Qy		
1544	Qy		
1545	Qy		
1546	Qy		
1547	Qy		
1548	Qy		
1549	Qy		
1550	Qy		
1551	Qy		
1552	Qy		
1553	Qy		
1554	Qy		
1555	Qy		
1556	Qy		
1557	Qy		
1558	Qy		
1559	Qy		
1560	Qy		
1561	Qy		
1562	Qy		
1563	Qy		
1564	Qy		
1565	Qy		
1566	Qy		
1567	Qy		
1568	Qy		
1569	Qy		
1570	Qy		
1571	Qy		
1572	Qy		
1573	Qy		
1574	Qy		
1575	Qy		
1576	Qy		
1577	Qy		
1578	Qy		
1579	Qy		
1580	Qy		
1581	Qy		
1582	Qy		
1583	Qy		
1584	Qy		
1585	Qy		
1586	Qy		
1587	Qy		
1588	Qy		
1589	Qy		
1590	Qy		
1591	Qy		
1592	Qy		
1593	Qy		
1594	Qy		
1595	Qy		
1596	Qy		
1597	Qy		
1598	Qy		
1599	Qy		
1600	Qy		
1601	Qy		
1602	Qy		
1603	Qy		
1604	Qy		
1605	Qy		
1606	Qy		
1607	Qy		
1608	Qy		
1609	Qy		
1610	Qy		
1611	Qy		
1612	Qy		
1613	Qy		
1614	Qy		
1615	Qy		
1616	Qy		
1617	Qy		
1618	Qy		
1619	Qy		
1620	Qy		
1621	Qy		
1622	Qy		
1623	Qy		
1624	Qy		
1625	Qy		
1626	Qy		
1627	Qy		
1628	Qy		
1629	Qy		
1630	Qy		
1631	Qy		
1632	Qy		
1633	Qy		
1634	Qy		
1635	Qy		
1636	Qy		
1637	Qy		
1638	Qy		
1639	Qy		
1640	Qy		
1641	Qy		
1642	Qy		
1643	Qy		
1644	Qy		
1645	Qy		
1646	Qy		
1647	Qy		
1648	Qy		
1649	Qy		
1650	Qy		
1651	Qy		
1652	Qy		
1653	Qy		
1654	Qy		
1655	Qy		
1656	Qy		
1657	Qy		
1658	Qy		
1659	Qy		
1660	Qy		
1661	Qy		
1662	Qy		
1663	Qy		
1664	Qy		
1665	Qy		
1666	Qy		

[illegible]

Db 3759 CAAACAACTTGTGGCTTCTCGTCGGAGGAGCAACTGCTTCGAGAACGACAGAT 3818  
Qy 3747 GGCACGCGCTCCCTTTGCGCTCTGTAAATGTGCGCTTGGAAACAGATGAGGAGCCTCCTGA 3806  
Db 3819 GGCACGCGCTCCCTTTGCTTCTGTGATGAGCCCTGGAAAGTAGGAGCTGAAACAGACAGA 3878  
Qy 3807 TCTTATTGGGGGAGTATAAGACTGTTCCCAAAACCATTTGCACTGGAGCCGTGTTTGG 3866  
Db 3879 CTTTCTGAGGGCCATTAGAGGGTCCCTTAAGCCAATAGCCATCGAACCCCTGGCTGG 3938  
Qy 3867 CAACAAAGCGCTGCTCTCTGTGTTTGTGAGGCTCCCTCGAGGCTTGGGTGGCATCCC 3926  
Db 3939 GAACAGAGCTGTGCTCCTGACTGTGTTTCTGTCTACCGAGAGGATCTTCAGGGCGCCC 3998  
Qy 3927 TCCTCCTGGCAGTCAGGTCTTGTGTCGGCCAGCGCCCTGCTGGACATTTCTCAGCAGAT 3986  
Db 3999 ACCCCCTGGGAGTCAGGCTTGTCTATCGCCAGTGCCTGTATAGATCTCAGAGAGAA 4058  
Qy 3987 GCGATAGTGTACAAGAGAAGTCAGGAGCCGTGAGAAACCGGAAGCAGACAGCCCCC 4043  
Db 4059 GCCTTCTGATATAAAGACAAGACTTCTGGAGTCCGCAATCGGAAGCACTCTCCAC 4115

Search completed: March 4, 2004, 18:09:10  
Job time : 971.587 secs

***This Page Blank (uspto)***

Result No.	Score	Query #	Length	DB	ID	Description
1	3988.4	98.0	8589	9	US-09-893-238-14	Sequence 14, Appl
2	3499.8	81.0	4072	9	US-09-893-238-16	Sequence 16, Appl
3	3326.4	85.8	8827	9	US-09-893-238-1	Sequence 1, Appl
4	2051.8	50.4	2625	9	US-09-893-238-18	Sequence 18, Appl
5	1853	45.6	6730	9	US-09-893-238-12	Sequence 12, Appl
6	1461.8	35.9	6373	16	US-10-197-824-1	Sequence 1, Appl
7	1451.8	35.7	2419	9	US-09-893-238-8	Sequence 8, Appl
8	463	11.4	1051	9	US-09-893-238-10	Sequence 10, Appl
C 9	277.4	6.8	625	14	US-10-198-846-5859	Sequence 5859, Ap
10	250.8	6.2	531	9	US-09-864-761-15900	Sequence 15900, A
C 11	246	6.0	246	9	US-09-736-457-966	Sequence 966, App
C 12	246	6.0	246	9	US-09-902-941-966	Sequence 966, App
C 13	246	6.0	246	9	US-09-849-626-966	Sequence 966, App
C 14	246	6.0	246	14	US-10-017-754-966	Sequence 966, App
C 15	246	6.0	246	14	US-10-113-872-966	Sequence 966, App



Db 2514 GCGAATAATGCACTCATCTCAGAGCATGTCCAGCTCACTTAAACCCCATGGGTGGGCT 2573  
Qy 2355 TCGGAAGATCAATGTGTCTCTACTGTGCTGGGAGATATGTCCCATTTACAAATAGTTT 2414  
Db 2574 TCGGAAGATCAATGTGTCTCTACTGTGCTGGGAGATATGTCCCATTTACAAATAGTTT 2633  
Qy 2415 ACTACAGTGCATGCGCTCTGAGCCAGTGCATGCTGATTTCTGTGGAAATTTATCAGAAC 2474  
Db 2634 ACTACAGTGCATGCGCTCTGAGCCAGTGCATGCTGATTTCTGTGGAAATTTATCAGAAC 2693  
Qy 2475 CAGTACTCGGGGACTGAAGGCTGCATCTGCAATCAACCCCACTCAATGGTAGTCTGTGA 2534  
Db 2694 CAGTACTCGGGGACTGAAGGCTGCATCTGCAATCAACCCCACTCAATGGTAGTCTGTGA 2753  
Qy 2535 AAGGCTGCAACCAACAGTGTGAAGTGCAGTGCAGGACACCATGTGCTGAGCAGCATG 2594  
Db 2754 AAGGCTGCAACCAACAGTGTGAAGTGCAGTGCAGGACACCATGTGCTGAGCAGCATG 2813  
Qy 2595 TGGAGATTGACACAGGCGAGCTCTGAGTGCATGTGGTGACAGCAACATGAAGCAGTGT 2654  
Db 2814 TGGAGATTGACACAGGCGAGCTCTGAGTGCATGTGGTGACAGCAACATGAAGCAGTGT 2873  
Qy 2655 GGAATCAATGCTATGTGGCTCTCTTTCCTGCTGAGTGTATGGAATGTATACGAT 2714  
Db 2874 GGAATCAATGCTATGTGGCTCTCTTTCCTGCTGAGTGTATGGAATGTATACGAT 2933  
Qy 2715 GAGCACTGCGCCCTGAAATTTGTTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGCA 2774  
Db 2934 GAGCACTGCGCCCTGAAATTTGTTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGCA 2993  
Qy 2775 ACCAGCTGTGGTGTGATCTGATCCAGCAATCTGCAAGGAAATGATAGAGG 2834  
Db 2994 ACCAGCTGTGGTGTGATCTGATCCAGCAATCTGCAAGGAAATGATAGAGG 3053  
Qy 2835 TTCTATTAAGGACCACTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGCA 2894  
Db 3054 TTCTATTAAGGACCACTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTGAGCA 3113  
Qy 2895 GCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCATCACTGCTTTTATTCATCA 2954  
Db 3114 GCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCATCACTGCTTTTATTCATCA 3173  
Qy 2955 CTGTCCAGCTTGCAATGMAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAA 3014  
Db 3174 CTGTCCAGCTTGCAATGMAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAA 3233  
Qy 3015 GTGTGAGAACTGACACAGGCAAGCACTGCGAGCTGCATATCTGGCTTCTACGGTGA 3074  
Db 3234 GTGTGAGAACTGACACAGGCAAGCACTGCGAGCTGCATATCTGGCTTCTACGGTGA 3293  
Qy 3075 TCCACCAATGGAGGAAATGTGAGCCATGCAAGTGCATGCGAGCAAGTGTCTGTGCAA 3134  
Db 3294 TCCACCAATGGAGGAAATGTGAGCCATGCAAGTGCATGCGAGCAAGTGTCTGTGCAA 3353  
Qy 3135 CACCAACAGGCAAGTGTCTGCAACCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAG 3194  
Db 3354 CACCAACAGGCAAGTGTCTGCAACCAAGGCTGCAAGGCTGCAAGGCTGCAAGGCTGCAAG 3413  
Qy 3195 ATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATATCTCT 3254  
Db 3414 ATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATATCTCT 3473  
Qy 3255 TCTTATGACTATCAGTTCACCTTGTATCTGCAAGGAAATGATGCTATATACAGC 3314  
Db 3474 TCTTATGACTATCAGTTCACCTTGTATCTGCAAGGAAATGATGCTATATACAGC 3533  
Qy 3315 TATCAATTTGTGGTACTCTGACGAAACAAACAGGAAATTTGGACATGTTCAATGCG 3374  
Db 3534 TATCAATTTGTGGTACTCTGACGAAACAAACAGGAAATTTGGACATGTTCAATGCG 3593  
Qy 3375 CTCGAAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTAGCTGGAACCCAGGC 3434  
Db 3594 CTCGAAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTAGCTGGAACCCAGGC 3653

Qy 3435 TGGAGAGAGATGCTGTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494  
Db 3654 TGGAGAGAGATGCTGTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713  
Qy 3495 TAAATGAGAAGTTTGAATTTTGGCAACCCCAAAATATCATTCTTTGTTTATGTCAGTAA 3554  
Db 3714 TAAATGAGAAGTTTGAATTTTGGCAACCCCAAAATATCATTCTTTTATGTCAGTAA 3773  
Qy 3555 TTTCACTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCACAGCAATTTTATGGACCT 3614  
Db 3774 TTTCACTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCACAGCAATTTTATGGACCT 3833  
Qy 3615 GGTACAGTTCCTGTCGACTTTCTTCACTGTTTCTCTCTCTTGTGCTGCTGGGCTGCTGT 3674  
Db 3834 GGTACAGTTCCTGTCGACTTTCTTCACTGTTTCTCTCTCTTGTGCTGCTGGGCTGCTGT 3893  
Qy 3675 GGTTTGGAAGATCAAAACAAAGTTGTGGGCTCTCAGACGCTAGAGAGCAACTTTCTTCGAGA 3734  
Db 3894 GGTTTGGAAGATCAAAACAAAGTTGTGGGCTCTCAGACGCTAGAGAGCAACTTTCTTCGAGA 3953  
Qy 3735 GATGCAACAGATGCGCAGCGCTCCCTTTGCTCTGTAAATGCTGCTTGGAAACAGATGA 3794  
Db 3954 GATGCAACAGATGCGCAGCGCTCCCTTTGCTCTGTAAATGCTGCTTGGAAACAGATGA 4013  
Qy 3795 GAGGCTCTCTGATCTTATTGGGGGAGTATAAAGACTGTTCCCAAAACCATTTGCACTGGA 3854  
Db 4014 GAGGCTCTCTGATCTTATTGGGGGAGTATAAAGACTGTTCCCAAAACCATTTGCACTGGA 4073  
Qy 3855 GCGGTGTTTGGCAACAAAGCGCTGCTCTCTCTGTGTTGTGAGGTCCCTCGAGGCT 3914  
Db 4074 GCGGTGTTTGGCAACAAAGCGCTGCTCTCTCTGTGTTGTGAGGTCCCTCGAGGCT 4133  
Qy 3915 GGTGGCATCTCCTCTCTGCGGAGTCAAGTCTGCTGCTGCGCCAGCGCTGCTGAGCAT 3974  
Db 4134 GGTGGCATCTCCTCTCTGCGGAGTCAAGTCTGCTGCTGCGCCAGCGCTGCTGAGCAT 4193  
Qy 3975 TTCTCAGCAGATGCGGATGATGTAAGGAACTCAGAGCGCTGAGAAACCGGAAGCA 4034  
Db 4194 TTCTCAGCAGATGCGGATGATGTAAGGAACTCAGAGCGCTGAGAAACCGGAAGCA 4253  
Qy 4035 GCAGCCCTCTCAGCAGCTGCGACCTGCACTGA 4068  
Db 4254 GCAGCCCTCTCAGCAGCTGCGACCTGCACTGA 4287

## RESULT 2

US-09-893-238-16  
; Sequence 16, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 4072  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-893-238-16  
Query Match 86.0%; Score 3499.8; DB 9; Length 4072;



Best Local Similarity 98.8%; Pred. No. 0; Matches 3525; Conservative 0; Mismatches 42; Indels 0; Gaps 0;									
Qy	15	GGCGGCAACTGAGCAAGCTGAGGAGGAGGAGCGCGCGGAGCGGAGCGCTCGCGGCGAG	74						
Db	234	GGAGGCGGAGCGCGCTGCGGTGCGCGCGCGGTGCTCGCGCTCGCGCGGAGCGGCGCA	293						
Qy	75	GAGGCGGCGCGCACTGTGTCAACGCGGTGCTGCAACCCCTGGCCACCGCGCGAGTGGCT	134						
Db	294	GGAATGTGACCGGCGGTGTCAACGCGGTGCTGCAACCCCTGGCCACCGCGCGAGTGGCT	353						
Qy	135	CTGCCCGCGCGGTGGGTGGCGGAGCAATGCCAGCACCTGCGGGCGCGCTTCAGACTAAC	194						
Db	354	CTGCCCGCGCGGTGGGTGGCGGAGCAATGCCAGCACCTGCGGGCGCGCTTCAGACTAAC	413						
Qy	195	TGATCTTCTGGGTGTGTGACAGTGGACCTGGAATTAATAACAACCAAGTGCAC	254						
Db	414	TGATCTTCTGGGTGTGTGACAGTGGACCTGGAATTAATAACAACCAAGTGCAC	473						
Qy	255	GTGGCTCATTGAAGGAGCAGCAATAGAAATAGAGACTTGTTCATCACTTTTGTCTAC	314						
Db	474	GTGGCTCATTGAAGGAGCAGCAATAGAAATAGAGACTTGTTCATCACTTTTGTCTAC	533						
Qy	315	AGAGTGTAGTGGGACCAATTAATATGTATGATGGGAGCTCAATTTATGACCGCTAGT	374						
Db	534	AGAGTGTAGTGGGACCAATTAATATGTATGATGGGAGCTCAATTTATGACCGCTAGT	593						
Qy	375	TGCTGCATTTAGTGGCTCAATGTCTGTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT	434						
Db	594	TGCTGCATTTAGTGGCTCAATGTCTGTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT	553						
Qy	435	TGTTGCCACATCAGGTTATGCCCTGTGCTGATTTTTTTAGTGTGCTGTATTAATTTGAC	494						
Db	654	TGTTGCCACATCAGGTTATGCCCTGTGCTGATTTTTTTAGTGTGCTGTATTAATTTGAC	713						
Qy	495	TGGATTTAATTTACTTACAGTTTGTATGTGTCGCAATTAATGCTCAGGCGGAGGACA	554						
Db	714	TGGATTTAATTTACTTACAGTTTGTATGTGTCGCAATTAATGCTCAGGCGGAGGACA	773						
Qy	555	GTGTAGATCAGTAATAGCAGCGATCTGTGTGAATGTGAATGTTCTGAAAACTGGAAAGG	614						
Db	774	GTGTAGATCAGTAATAGCAGCGATCTGTGTGAATGTGAATGTTCTGAAAACTGGAAAGG	833						
Qy	615	TGAAGCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGCACTG	674						
Db	834	TGAAGCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGCACTG	893						
Qy	675	CAATTCAAGTGATGTACAGAGATGCTCTGCTTCTCAGACTGCGAGGCTCCTGGATGTTT	734						
Db	894	CAATTCAAGTGATGTACAGAGATGCTCTGCTTCTCAGACTGCGAGGCTCCTGGATGTTT	953						
Qy	735	AGTTCTGTACCAAGTAAACAGTCAATTTTGGATCTCGAGAGGAATATTCTAACTTAAAGCT	794						
Db	954	AGTTCTGTACCAAGTAAACAGTCAATTTTGGATCTCGAGAGGAATATTCTAACTTAAAGCT	1013						
Qy	795	CCCAGAGCATCTCATAAAGCTGTGTCATGAGAAACATTTATGTGGTTGTTGGAGGATA	854						
Db	1014	CCCAGAGCATCTCATAAAGCTGTGTCATGAGAAACATTTATGTGGTTGTTGGAGGATA	1073						
Qy	855	TATGTTCAACCACTCAGATTTAACAATGTTTCTAGCGTATGACCTTGTCTAGGGAGTG	914						
Db	1074	TATGTTCAACCACTCAGATTTAACAATGTTTCTAGCGTATGACCTTGTCTAGGGAGTG	1133						
Qy	915	GCITTCAGCTAAACCGTTCCTGACAGATGTGGTTGTGTAGATATGTCATCTTTTGGCAAT	974						
Db	1134	GCITTCAGCTAAACCGTTCCTGACAGATGTGGTTGTGTAGATATGTCATCTTTTGGCAAT	1193						
Qy	975	ATACAAGGATAAATTTTACATGTATGAGGAGAAAAATTTGATTTCACTGGGAATGTGACCAA	1034						
Db	1194	ATACAAGGATAAATTTTACATGTATGAGGAGAAAAATTTGATTTCACTGGGAATGTGACCAA	1253						
Qy	1035	TGAGTTGAGAGTTTTTACATTAATAGTATGAGGTTGTTGTCACCCCTTAAGGCAAA	1094						

Db	1254	TGAGTTGAGAGTTTTTTCACATTCATAATAGTATGAGTCAATGGGTGTTGTTGACCCCTTAAGGCAAA	1313						
Qy	1095	GGAGCATGTGACAGTGGTTGGGCACTCTGACACATTTGTACACTGAAGAAATGGCGAGT	1154						
Db	1314	GGAGCATGTGACAGTGGTTGGGCACTCTGACACATTTGTACACTGAAGAAATGGCGAGT	1373						
Qy	1155	GGTCATGCTGCTCATCTTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGACGGA	1214						
Db	1374	GGTCATGCTGCTCATCTTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGACGGA	1433						
Qy	1215	ATATGATTTGGATGAAGAACACATGGAGTATATACACACCCAGGCTGCCCTTGTGCAAGG	1274						
Db	1434	ATATGATTTGGATGAAGAACACATGGAGTATATACACACCCAGGCTGCCCTTGTGCAAGG	1493						
Qy	1275	GGTTTACGGCCATAGCAGTGTTCACGACCATAGGACAGGCGCCCTATACGTTTCAATGCTGG	1334						
Db	1494	GGTTTACGGCCATAGCAGTGTTCACGACCATAGGACAGGCGCCCTATACGTTTCAATGCTGG	1553						
Qy	1335	CTACAAGGCTTTTCAAGTGAAGTACCGGCTTGGAGTGAATCTCTACCGATATGATGT	1394						
Db	1554	CTACAAGGCTTTTCAAGTGAAGTACCGGCTTGGAGTGAATCTCTACCGATATGATGT	1613						
Qy	1395	GGATACCCAGATGTGGACCACTTCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGC	1454						
Db	1614	GGATACCCAGATGTGGACCACTTCTTAAGGACAGCCGATTTTCCGTTTACTTGCACACAGC	1673						
Qy	1455	TGTGATGTGAGTGGAAACCAATGCTGTGTGTTGGGGGAAAAACACACAATGACACATCTAT	1514						
Db	1674	TGTGATGTGAGTGGAAACCAATGCTGTGTGTTGGGGGAAAAACACACAATGACACATCTAT	1733						
Qy	1515	GAGCCATGGCGCCAAATGCTTCTTCAGATTTTCATGCTCATGACATTCGCTGTGACCG	1574						
Db	1734	GAGCCATGGCGCCAAATGCTTCTTCAGATTTTCATGCTCATGACATTCGCTGTGACCG	1793						
Qy	1575	CTGCTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC	1634						
Db	1794	CTGCTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC	1853						
Qy	1635	AGTCTTACAACAGACCACTGATGTGTCGGTGTTCATAGTCTCTCCCTCAGCGA	1694						
Db	1854	AGTCTTACAACAGACCACTGATGTGTCGGTGTTCATAGTCTCTCCCTCAGCGA	1913						
Qy	1695	CATCTCTGATTTTCACTCTCGGAACTGTGATGCGCATCGGAGTGAAGCCGCTTGTATGC	1754						
Db	1914	CATCTCTGATTTTCACTCTCGGAACTGTGATGCGCATCGGAGTGAAGCCGCTTGTATGC	1973						
Qy	1755	AGCAGGACCTGGATTTGCTGTGTGTGGAACAACAGGCTGCTCAGTGTATCTGTGGGC	1814						
Db	1974	AGCAGGACCTGGATTTGCTGTGTGTGGAACAACAGGCTGCTCAGTGTATCTGTGGGC	2033						
Qy	1815	GCTGGCACTGATCAACAAGAAAGAAAGTTAAATCAGAATGTTTTTCCAAAAGAACTCT	1874						
Db	2034	GCTGGCACTGATCAACAAGAAAGTTAAATCAGAATGTTTTTCCAAAAGAACTCT	2093						
Qy	1875	TGACCATGACAGATGTGACCAAGCAACAGATTTGTTACAGTGTGCAACCAACCAATGA	1934						
Db	2094	TGACCATGACAGATGTGACCAAGCAACAGATTTGTTACAGTGTGCAACCAACCAATGA	2153						
Qy	1935	CTGGCACTGGTGCATGACCATTTGCTGCCAGGAAACACAGCTGCTCAGAAGCCAGAT	1994						
Db	2154	CTGGCACTGGTGCATGACCATTTGCTGCCAGGAAACACAGCTGCTCAGAAGCCAGAT	2213						
Qy	1995	CTCCATTTTATAGTATGAGAAATTCGCCCAAGGATAACCTATGTPACTCTGTAAACAGAA	2054						
Db	2214	CTCCATTTTATAGTATGAGAAATTCGCCCAAGGATAACCTATGTPACTCTGTAAACAGAA	2273						
Qy	2055	GACAGCTGACGAGCTGTGCGCTGGAGCAACCTGCGAGTGGGAGCCCGGATCAGGA	2114						
Db	2274	GACAGCTGACGAGCTGTGCGCTGGAGCAACCTGCGAGTGGGAGCCCGGATCAGGA	2333						
Qy	2115	GTGANTTCCCTGCGCAAAATATCTGTGGCATTTGCTGGCATTTGGTATGGAATCTCATG	2174						
Db	2334	GTGANTTCCCTGCGCAAAATATCTGTGGCATTTGCTGGCATTTGGTATGGAATCTCATG	2393						

```
QY 2175 TTTGAAATTTACTGCTGCAAGAGAAATATGACAAATGCTAAATGTTCTGTAGAACCA 2234
Db 2194 TTTGAAATTTACTGCTGCAAGAGAAATATGACAAATGCTAAATGTTCTGTAGAACCA 2453
QY 2235 CAATGCCCTTTGGCTTCTCTTCAACCCAGAGAGGTAGAAATTTGCTCTTAAGAGCT 2294
Db 2454 CAATGCCCTTTGGCTTCTCTTCAACCCAGAGAGGTAGAAATTTGCTCTTAAGAGCT 2513
QY 2295 GCGAATAATGCACTCATCTCAGAGCATGTCGAAGCTCACTTAAACCCCATGGGTGGGCT 2354
Db 2514 GCGAATAATGCACTCATCTCAGAGCATGTCGAAGCTCACTTAAACCCCATGGGTGGGCT 2573
QY 2355 TCGGAAGATCAATGTTGCTCTTCTGCTGCTGGAAGATATGTCCTCATTTACAAATAGTTT 2414
Db 2574 TCGGAAGATCAATGTTGCTCTTCTGCTGCTGGAAGATATGTCCTCATTTACAAATAGTTT 2633
QY 2415 ACTACAGTGGATGCGCTGAGCCAGTATGCTGGATCTCTGGAATTTTATCAGAAC 2474
Db 2634 ACTACAGTGGATGCGCTGAGCCAGTATGCTGGATCTCTGGAATTTTATCAGAAC 2693
QY 2475 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTA 2534
Db 2694 CAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTA 2753
QY 2535 AAGGCTGCAAAACCAAGTGTCAAGCAGTCCGAGACCATGTGCTTGAGGACAGCATG 2594
Db 2754 AAGGCTGCAAAACCAAGTGTCAAGCAGTCCGAGACCATGTGCTTGAGGACAGCATG 2813
QY 2595 TGGAGATTGACACAGCGGAGCTCTGAGTCCATGTGCTGAGCAACATCAAGCAGTGTGT 2654
Db 2814 TGGAGATTGACACAGCGGAGCTCTGAGTCCATGTGCTGAGCAACATCAAGCAGTGTGT 2873
QY 2655 GGAATCAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTATGGAATGTATACGAT 2714
Db 2874 GGAATCAATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTATGGAATGTATACGAT 2933
QY 2715 GAGCACTGCCCCCTGAAATTTGTTGAGGCTACTGTACTGTGAGTCTGAGTCTGAGCA 2774
Db 2934 GAGCACTGCCCCCTGAAATTTGTTGAGGCTACTGTACTGTGAGTCTGAGTCTGAGCA 2993
QY 2775 ACCAGCTGTGGCTGTACTGATCCAGCAATATCTGCAAGGAAATGATAGAGG 2834
Db 2994 ACCAGCTGTGGCTGTACTGATCCAGCAATATCTGCAAGGAAATGATAGAGG 3053
QY 2835 TTCTATAAGNACAGTGAAGTGCCTTGGAGCCCTTACAGGAATTTCTATCCACA 2894
Db 3054 TTCTATAAGNACAGTGAAGTGCCTTGGAGCCCTTACAGGAATTTCTATCCACA 3113
QY 2895 GCGCCCTGCTCAATTCAGCATGTGCTAGAGGACAGCATACAACTGGTCTTTTCAATCA 2954
Db 3114 GCGCCCTGCTCAATTCAGCATGTGCTAGAGGACAGCATACAACTGGTCTTTTCAATCA 3173
QY 2955 CTGTCCAGCTTGGCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTAGAA 3014
Db 3174 CTGTCCAGCTTGGCAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTAGAA 3233
QY 3015 GTGTGAGAACTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGTCTTACGGTGA 3074
Db 3234 GTGTGAGAACTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGTCTTACGGTGA 3293
QY 3075 TCCCAACAATGGAGGAAATGTGAGCCATGCAAGTGCATGGGACGGCTCTGTGCA 3134
Db 3294 TCCCAACAATGGAGGAAATGTGAGCCATGCAAGTGCATGGGACGGCTCTGTGCA 3353
QY 3135 CACCAACAGGGCAAGTGTCTGACACCAAGGCGTCAAGGGGAGGAGTGGCAGCT 3194
Db 3354 CACCAACAGGGCAAGTGTCTGACACCAAGGCGTCAAGGGGAGGAGTGGCAGCT 3413
QY 3195 ATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATATCTCT 3254
Db 3414 ATGTGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAAACATGTTATATCTCT 3473
```

```
QY 3255 TCTTATTGACTATCAGTTTCACTTTAGTCTATATCCAGAGAGATGATCGCTATTACAGC 3314
Db 3474 TCTTATTGACTATCAGTTTCACTTTAGTCTATATCCAGAGAGATGATCGCTATTACAGC 3533
QY 3315 TATCAATTTTGTGCTACTCTGACGAAACAAACAGAGATTTGGACATGTTTCATCAATGC 3374
Db 3534 TATCAATTTTGTGCTACTCTGACGAAACAAACAGAGATTTGGACATGTTTCATCAATGC 3593
QY 3375 CTCCAAGAAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGC 3434
Db 3594 CTCCAAGAAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGC 3653
QY 3435 TGGAGAGAGATGCTGTTTCAAAAACCAACATTAAGAGTACAAAGATGTTTCTC 3494
Db 3654 TGGAGAGAGATGCTGTTTCAAAAACCAACATTAAGAGTACAAAGATGTTTCTC 3713
QY 3495 TAATGAGAAGTTTGATTTTGGCAACCAACCAATATCACTTTCTTGTATGTCAGTAA 3554
Db 3714 TAATGAGAAGTTTGATTTTGGCAACCAACCAATATCACTTTCTTGTATGTCAGTAA 3773
QY 3555 TTTCACTGGCCCATCAAAATTCAGAT 3581
Db 3774 TTTCACTGGCCCATCAAAATTCAGAT 3800

RESULT 3
US-09-893-238-1
; Sequence 1, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-1

Query Match      81.8%; Score 3326.4; DB 9; Length 8827;
Best Local Similarity 91.1%; Pred. No. 0;
Matches 3534; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 189 ACTAACTGGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTAATAACAAACGAA 248
Db 677 ACTAACTGGCTCTTCTGGATTTGTAAACAGATGGACCTGGGAATTAATAACAAACGAA 736
QY 249 GTGCACGTGGCTCAATGAGGACAGCAAAATAGAAATAGAGCTTCGTTTCATCATTT 308
Db 737 GTGCACATGGCTCAATGAGGACAGCAAAATAGAAATAGAGCTTCGTTTCATCATTT 796
QY 309 TGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACCTCAATTTATGCACC 368
Db 797 TGCTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACCTCAATTTATGCACC 856
QY 369 GCTAGTTGCTGCATTTTGTGCTCATTTGTTCCTGAGAGAGATGGCAATGAGACTGCC 428
Db 857 TCTGATTTGCTGCTTTTGTGCTCATTTGTTCCTGAAAGAGAGATGGCAATGAGACTGCC 916
QY 429 TGAGGTTGTTGCCACATCAGGTTATGCTTGTCTCATTTTATGATGATGCTGCTTATAA 488
```

Db 917 TGAGGTCACCTGTCACCTTCAGGTTATGCACTGCTGCAATTTTTCAGTGTGCTGTATATAA 976  
Qy 489 TTTGACTGGATTTTAATATTACTTACAGTTTGTATATGTCCTCAATAATCTGCTCAGGCCG 548  
Db 977 TCTGACTGGATTTAATATCACTTACAAATTTTGACATGTGTCCGAATAATTGCTCAGGCCG 1036  
Qy 549 AGGAGAGTGAAGATCAGTAATAAGCAGCGATACCTGTTGAATGTGAATTTCTGAAAACCTG 608  
Db 1037 AGGAGAGTGAAGCAGTAAACAGCAGCAGCGCTGTTGAGTGTGAATTTCTGAAAACCTG 1096  
Qy 609 GAAAGGTGAAGATGTGACATTTCTCTCATGTACAGACAACTGTGTTTCTCTCATCGAGG 668  
Db 1097 GAAAGGGGAGTCGTGTGACATTTCTCTCATGTACAGACAACTGTGTTTCTCTCATCGAGG 1156  
Qy 669 CATCTGCAATTTCAAGTGTGTCAGAGATGCTCTCTCTCTCATGACTGCGAGGTCCTGG 728  
Db 1157 CATCTGTAATGCAAGCGATACAGAGGGTGTCTCTCTCTCTCATGCTGCGAGGTCCTGG 1216  
Qy 729 ATGTTCAAGTTCCTGTACAGCTAACCAAGTCAATTTTGGACTCGAGAGGAATATTCTAACTT 788  
Db 1217 ATGTTCAATTTCTGTGCCAGTAAACAGTCTTTTGGACTCGAGAGGAATATTCTGAATTT 1276  
Qy 789 AAGCTCCCGAGACCTCATAAAGCTGTGGTCAATGGAAACATTTATGTGGTGTGGTGG 848  
Db 1277 AAGCTTTCCAGAGCCCTCTCATAAAGCTGTGGTCAATGGAAATATTAATGTGGTGTGGTGG 1336  
Qy 849 AGGATATATGTTCAACCACTCAGATTATTAACATGTTTCTAGCGTATGACCTTGTCTTAG 908  
Db 1337 CGGATATATGTTCAACCACTCAGATTATTAACATGTTTCTAGCGTATGACCTGACTTCTAG 1396  
Qy 909 GAGTGGCTTCCATTAACCGTTCTGTGAACATGTGGTGTGTAGATATGTTCTATCTTTT 968  
Db 1397 GGAATGGCTTCCATTAACCACTTCTGTGAACAGTGTGGTGTGTAGATATGTTCTATCTTT 1456  
Qy 969 GGCATTATACAAGGATAAATTTACATGTATGAGGAGAAATTTGATTTCACTGGGAATGT 1028  
Db 1457 GGCATTATACAAGGATAAATTTACATGTATGAGGAGAAATTTGATTTCAACAGGGAATGT 1516  
Qy 1029 GACCAATGAGTGTGAGAGTGTGTTTCAATTTCAATGTAATCATGCGGTATTTGTTAACTCCGAA 1576  
Db 1517 GACCAATGAGTGTGAGAGTGTGTTTCAATTTCAATGTAATCATGCGGTATTTGTTAACTCCGAA 1576  
Qy 1089 GGCAGAGGACGATGACGAGTGTGGGCACTGTGACACATTTGTACACTGAAGAATGG 1148  
Db 1577 AGCTAAGGATCAGTATGCGATGTGTGACACTGACACACATTTGTACTGGCATCTGG 1636  
Qy 1149 CCGAGTGGTCATGCTGGTCACTTTTGGTCACTGCTCCCTCTCTATGAGATATATAAGCAATGT 1208  
Db 1637 CCGTGTGGTCACTGTTGGTCACTTCCGTCACTGCTCCCACTCTATGAGATATATAAGCGTTGT 1696  
Qy 1209 GCAGGATATGATTTGGATGAAGACACATGGAGTATATTACACCCAGGCTGCTGCTGTG 1268  
Db 1697 GCAGGATATGATTTGGATGAAGACACATGGAGTATATTACATCTCAGGGTGTCTGTTG 1756  
Qy 1269 GCAAGGGGGTTAGCGGCATAGCAGTGTATTACGACCATAGACAGCGGCCCTTATACGTTCA 1328  
Db 1757 GCAAGGGGGTTAGCGGCATAGTGTATTATGATGACAGCAAGGCTCTGTACGTTCA 1816  
Qy 1329 TGGTGGCTACAAGGCTTTAGTGGCAATAGATACCGGCTTGGAGATGATCTCTACCGATA 1388  
Db 1817 TGGTGGCTACAAGGCTTTAGTGGCAATAGATACCGGCTTGGAGATGATCTCTACAGATA 1876  
Qy 1389 TGAATGGATACCCAGATGTGGACCAATTTTAAGGACAGCCGATTTTTCGTTACTTTGCA 1448  
Db 1877 CGATGTGATCTCAGATGTGGACCAATTTTAAGGACAGCCGATTTTTCGTTACTTTGCA 1936  
Qy 1449 CACAGCTGTGATGTAGTGGGAACCAATGCTGTGGTGTGTTGGGGGAAACACACATGACAC 1508  
Db 1937 TACAGCTGTGTATGTAGTGGGAACCAATGCTGTGGTGTGTTGGGGGAAACACACATGACAC 1996  
Qy 1509 ATCTATGAGCCATGGGCCAAATGCTTCTCTCATGATTTTCTGAGATTTTCTGCGCTATGATGCTGCTG 1568  
Db 1997 TTCCATGAGCACGGTGCCAAATGCTTCTCTCTCGGACTTCATGGCTTATGACATGCTGTTG 2056

Qy 1569 TGACCGCTGGTCACTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCCA 1628  
Db 2057 TGACCGATGGTCACTGCTTCCAGACCTGAGCTCCATCATGATGTCAACAGATTTGGCCCA 2116  
Qy 1629 TTCAGCAGTCTTTACAAACAGCAGCATGTATGTGTTCCGGTGGTGTTCATAGTCTCTCTCT 1688  
Db 2117 TTCAGCAGTCTTTTACAAACAGCAGCATGTATGTGTTCCGGGCGCTTCAACAGCCTCTCTCT 2176  
Qy 1689 CAGCGCATCTCTGATTTTCACTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTTG 1748  
Db 2177 CAGTGAAGTCTTTGCTTTTACCTCGGAGCAGTGGCATGCAACCGCAGTGAAGCTGCTTG 2236  
Qy 1749 TTTAGCAGCAGGACCTGATTTCCGTTGTGTGTGGAACACAGGGTGTCTCAGTGTATCTC 1808  
Db 2237 TGTGGCAGCAGGACCTGATTTCCGTTGTGTGTGGAACACACAGTGTCTCGATGTACTCTC 2296  
Qy 1809 GTGGGCGCTGGCACTGATGNAACAGAGAAAGTTAAATCAGAAATGTTTTCCTAAAG 1868  
Db 2297 CTGGGAGTGGCAATCTGAAGAACAGCAGAAAGTTAAATCAGAGTGTCTTTCTTAAAG 2356  
Qy 1869 AACTCTTTCACCATGACAGATGTGACCAAGCAGACACAGATTGTTTACAGCTGCAAGCCAAAC 1928  
Db 2357 AACCTTTGACCATGACAGATGTGACCAAGCAGACACAGATTGTTTACAGCTGCAAGCCAAATC 2416  
Qy 1929 CAATGACTGCCACTGTGGTGCATGACCAATTTGTGTCTCCAGGAACACAGCTGCTCAGNAGG 1988  
Db 2417 CAATGACTGCCACTGTGGTGCATGACCAATTTGTGTCTCCCTGTGAAACACAGCTGCAAGNAGG 2476  
Qy 1989 CCAGATCTCCATTTTAAAGTATGAGAAATTTGCAAGTGTGCAAGGATAAAGCTATGTACTGTAA 2048  
Db 2477 CCAGATCTCCATTTGCAAGTATGAGATTTGCCCAAGGATAAAGCTATGTACTGTAA 2536  
Qy 2049 CAAGAGACAGCTGTGAGAGCTGTGCTGAGCAGACAGAACTGCGAGTGGAGGCCCGGAA 2108  
Db 2537 TAAGAAACACAGCTGTGAGAGCTGTGCTGAGCAGACAGAACTGCGAGTGGAGGCCCGGAA 2596  
Qy 2109 TCAGAGTGCATGCTGCTGCGGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGAAA 2168  
Db 2597 TCAAGAGTGCATGCTGCTGCGGAAATATCTGTGGCAATTTGGCTGGCAATTTGGTTGAAA 2656  
Qy 2169 CTCATGTTTGAATAATTAATCTGCAAGAGGAATTTATGACAAATGCTAAATTTGTTGTAG 2228  
Db 2657 CTCGTTGTGAAATCACTACTGCTAAGGAGAAATTTATGACAAATGCTAAATTTGCTGTAG 2716  
Qy 2229 GAAACCAATGCTGCTTGTGCTTCTTACAAACCAAGAGGTAGAAATTTGTCCTTAA 2288  
Db 2717 GAACCAATGCTGCTTGTGCTTCTTACATCCCAAGAGAGGTAGAAATTTGTCCTTAA 2776  
Qy 2289 GCAGTGCAGATAATGCACTCATCTCAGAGCATGTCCAAGCTCACCTTAAACCCCATGGT 2348  
Db 2777 GCAGTTCGATTAATGCAATCATCTCAAGATGATGCTCCAAGCTCACTGACTCCATGGT 2836  
Qy 2349 CGGCTTTGCGAAGATCAATGTGCTTCTTACAAACCAAGAGATATGTCCTCATTTACAA 2408  
Db 2837 TSGTCTTGGAGATCAATGTGCTTCTTACTGCTGCGAGGATATGCTCCATTTACAAA 2896  
Qy 2409 TAGTTTACTACAGTGGATGCGCTGAGCCAGTGTGCTGAGATTTCTGTGGAAATTTTATC 2468  
Db 2897 TAGTTTGTGAGTGGATGCCATCTGAGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2956  
Qy 2469 AGAACCCAGTACTCGGGGACTGAAGGCTGCACCTGCAATCAACCCACTCAATGGTAGTGT 2528  
Db 2957 AGAGCTAGTACTCGGGGATTTAAAGGCTGCAACCTGCAATCAACCCCTCTCAATGGCAGCT 3016  
Qy 2529 CTGTGAAGGCTCTCAAAACCAAGTGTGAAGCAGTGGCGGACCACTGCTGCTTGGAGC 2588  
Db 3017 CTGTGAAGGCTCTCAAAACCAAGTGTGAAGCAGTGGCGGACCACTGCTGCTTGGAGC 3076  
Qy 2589 AGCATGTGAGATGTCAACAGCGGCACTGTGAGTGTGATGCTGCTGAGCAACATGAGACA 2648  
Db 3077 AGCGTGTGGGAGTGTCAACAGCAGCTGAGTGTGATGCTGCTGAGTGTGATGCTGAGTGTGAGCA 3136

QY 2649 GTGTGTGGAATCAATGCTATGTGGCTCTCTTCCCTTTTGCCAGTGTATGGAATGTA 2708  
DB GTGTGTGGAATCAATGCTATGTGGCTCTCTTCCCTTTTGCCAGTGTATGGAATGTA 3196  
QY 2709 TAGCATGAGCACTGCCCCCTGAAATTTGTCAGCTACTGTACTGTAGTCAATGCTT 2768  
DB TAGCATGAGCACTGCCCCCTGAAATTTGTCAGCTACTGTACTGTAGTCAATGCTT 3256  
QY 2769 GGAAGCAACCAAGCTGTGGCTGTGTACTGATCCAGCAATATGGAAGGGAATGCAAT 2828  
DB GGAAGCAACCAAGCTGTGGCTGTGTACTGATCCAGCAATATGGAAGGGAATGCAAT 3316  
QY 2829 AGAGGCTTCTTAAGAGGACAGTGAAGTGCCTTCGCAAGCCCTCAGAGAAATTTCTA 2888  
DB TGAGGCACTATAAGAGGACCTGTGAAGATGCCGTTCAGAGGCTCTGCAAGAAATTTGTA 3376  
QY 2889 TCCACAGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGAGATCAATCTGTCTTT 2948  
DB TCCACAGCCCTCTCTCAATTCAGCATGTGTCTAGAGGACAGAGATCAATCTGTCTTT 3436  
QY 2949 CATTCATGTCAGCTTGCATGCAACCGGCACAGTAATGATCAATCAGAGCATCTG 3008  
DB CATTCATGTCAGCTTGCATGCAACCGGCACAGTAATGATCAATCAGAGCATCTG 3496  
QY 3009 TGAGAGTGTGAGAACTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTCTA 3068  
DB TGAGAGTGTGAGAACTGACACAGGCAAGCACTGCGAGACCTGCATATCTGGCTCTA 3556  
QY 3069 CGGTGATCCCAATGAGGGAATGTGAGCCATGCAAGTCAATGGGACGCTCTCT 3128  
DB TGGTGACCCGACTAATGAGGGAATGTGAGCCATGCAAGTCAATGGGACGCTCTCT 3616  
QY 3129 GTGCAACACACAGGCAAGTGTCTGCAACCAAGGCGGTCAAGGGGAGAGTG 3188  
DB GTGCAACACACAGGCAAGTGTCTGCAACCAAGGCGGTCAAGGGGAGAGTG 3676  
QY 3189 CAGCTATGTGAGTGAAGAAATGATACCAAGAAACCTCTCAGAGGAACATGTTATTA 3248  
DB CAGCTATGTGAGTGAAGAAATGATACCAAGAAACCTCTCAGAGGAACATGTTATTA 3736  
QY 3249 TACTCTTCTATTGATATCAGTTCACCTTTAGTCTATCCAGGAGAGATGATCGTATTA 3308  
DB TACTCTTCTATTGATATCAGTTCACCTTTAGTCTATCCAGGAGAGATGATCGTATTA 3796  
QY 3309 CACAGCTATCAATTTTGTGCTACTCTCAGCAAAACAAACAGGATTTGACATGTCAT 3368  
DB CACAGCTATCAATTTTGTGCTACTCTCAGCAAAACAAACAGGATTTGACATGTCAT 3856  
QY 3369 CAATGCTTCAAGAAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAC 3428  
DB CAATGCTTCAAGAAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAC 3916  
QY 3429 CCAGGCTGAGAGAGATGCTGTTGTTTCAAAACCAACATTAAGGATCAAGAGATAG 3488  
DB CCAGGCTGAGAGAGATGCTGTTGTTTCAAAACCAACATTAAGGATCAAGAGATAG 3976  
QY 3489 TTTTCTTAATGAAAGTTTGAATTTTCGCAACCAACCAATATCACTTTCTTTGTTATGT 3548  
DB TTTTCTTAATGAAAGTTTGAATTTTCGCAACCAACCAATATCACTTTCTTTGTTATGT 4036  
QY 3549 CAGTAATTTTCACTGGCCCATCAAAATTCAGATGCTTCTCTCAGCAGCAAGAAATTTAT 3608  
DB CAGTAATTTTCACTGGCCCATCAAAATTCAGATGCTTCTCTCAGCAGCAAGAAATTTAT 4096  
QY 3609 GGAACCTGTGACAGTTCTCTGTAATTTTTCAGTGTGTTTCTCTCTCTCTCTCTGTCG 3668  
DB GGAACCTGTGACAGTTCTCTGTAATTTTTCAGTGTGTTTCTCTCTCTCTCTCTGTCG 4156  
QY 3669 TGTGTGTTTGAAGATCAAAAGTTTGTGGCCCTCCAGACGCTAGAGAGCAATCTCT 3728  
DB TGTGTGTTTGAAGATCAAAAGTTTGTGGCCCTCCAGACGCTAGAGAGCAATCTCT 4216  
QY 3729 TCGAGAGATGCAACAGATGGCCAGCGCTGCTCTGCTCTGTAATGTCGCTTGAAC 3788

DB 4217 TCGGAGATGCAACAGATGGCCAGCCCTTGTCTTCTGTAACCGTTGCTTGGAAAC 4276  
QY 3789 AGATGAGAGCTCTCTCATCTTATTGGGGGAGTATAAGACTGTTCCCAACCCATTGC 3848  
DB 4277 AGATGAGAGCTCTCTCATCTTATTGGGGGAGTATAAGACTGTTCCCAACCCATTGC 4336  
QY 3849 ACTGAGCCGCTGTTTGGCAACAAAGCCGCTGCTCTCTCTGTTTGTGAGGCTCCCTCG 3908  
DB 4337 ACTGAGCCGCTGTTTGGCAACAAAGCCGCTGCTCTCTCTGTTTGTGAGGCTCCCTCG 4396  
QY 3909 AGGCTGGTGGCATCCCTCTCTGGCAGTCAAGTCTTGTGTCGCGCAGCGCCCTGGT 3968  
DB 4397 AGGCTGGTGGCATCCCTCTCTGGCAGTCAAGTCTTGTGTCGCGCAGCGCCCTGGT 4456  
QY 3969 GGACATTTCTCAGCAGATGCCATAGTGTACAAGGAAAGTCAAGAGCCGCTGAGAAACCG 4028  
DB 4457 GGACATTTCTCAGCAGATGCCATAGTGTACAAGGAAAGTCAAGAGCCGCTGAGAAACCG 4516  
QY 4029 GAAGCAGAGCCCTCTGACAGCCTGGACCTGCATCTGA 4068  
DB 4517 GAAGCAGAGCCCTCTGACAGCCTGGACCTGCATCTGA 4556  
  
RESULT 4  
US-09-893-238-18  
; Sequence 18, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR FILING DATE: 09/245,041  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-893-238-18  
  
Query Match 50.4%; Score 2051.8; DB 9; Length 2625;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 2089; Conservative 0; Mismatches 62; Indels 0; Gaps 0;  
  
QY 15 GCGCGCAACTGAGCAAGGTGTAGGAGAGACGCGCGGCGGTGTCGGCTCGCGCCGAGGCGAA 293  
DB 234 GAGCGCGAGCCCTCGGTGCGCGGCGGTGTCGGCTCGCGCCGAGGCGAA 293  
QY 75 GAGCGCGGCGCGCACTGTGTCAACGCGGTGCGTCAACCCCTGGCAACCGCCAGTGGT 134  
DB 294 GGAATGTACCGCGGTGTGTCAACGCGGTGCGTCAACCCCTGGCAACCGCCAGTGGT 353  
QY 135 CTGCGCCCGCGGTGGGTGGCGCAATGCGAGCACTGCGGGGGCGCGCTTCAGACTAAC 194  
DB 354 CTGCGCCCGCGGTGGGTGGCGCAATGCGAGCAATGCCAGCACTGCGGGGGCGCGCTTCAGACTAAC 413  
QY 195 TGGATCTTCTGGGTGTGTGACAGATGACCTGGAAATATATAATACAAAGAGTGCAC 254  
DB 414 TGGATCTTCTGGGTGTGTGACAGATGACCTGGAAATATATAATACAAAGAGTGCAC 473  
QY 255 GTGCTCATTTGAAGCAGCAAGCAATAGATAATAGACTTCTGTTTCAATCATTTTGTCTAC 314  
DB 474 GTGCTCATTTGAAGCAGCAAGCAATAGATAATAGACTTCTGTTTCAATCATTTTGTCTAC 533



```
; SEQ ID NO 12
; LENGTH: 6370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-12

Query Match      45.6%; Score 1853; DB 9; Length 6370;
Best Local Similarity 96.3%; Pred. No. 0;
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;

QY 2035 ATGTACTACTGTAAACAAAGAACCAAGCTGCAGAGCTGTGCCCTGGACCCAGAACTGCCAG 2094
Db 1 ATGTACTACTGTAAACAAAGAACCAAGCTGCAGAGCTGTGCCCTGGACCCAGAACTGCCAG 60

QY 2095 TGGAGCCCCGGAATCAGAGTGCATTTGCCCTGCCGGAATAATCTGTGGCAATGGCTGG 2154
Db 61 TGGAGCCCCGGAATCAGAGTGCATTTGCCCTGCCGGAATAATCTGTGGCAATGGCTGG 120

QY 2155 CATTTGGTTGGAATCATGTTTGAATAATTACTTGCACAGGAGAAATTATGACAATGCT 2214
Db 121 CATTTGGTTGGAATCATGTTTGAATAATTACTTGCACAGGAGAAATTATGACAATGCT 180

QY 2215 AAATTGTTCTGTAGGAACCAATGCCCTTTTGGCTTCTTTACAAACCCAGAAAGGTA 2274
Db 181 AAATTGTTCTGTAGGAACCAATGCCCTTTTGGCTTCTTTACAAACCCAGAAAGGTA 240

QY 2275 GAATTGTCCTTAAGCAGCTGGAATATGAGTCACTCTCAGAGCATGTCGAGCTCACC 2334
Db 241 GAATTGTCCTTAAGCAGCTGGAATATGAGTCACTCTCAGAGCATGTCGAGCTCACC 300

QY 2335 TTAACCCCATGGGTGGCCCTT---CGGAAGATCAATGTCTCTACTGG-----TGCTGGG 2386
Db 301 TTAACCCCATGGGTGGCCCTT---CGGAAGATCAATGTCTCTACTGGGTCTGGGGAAG 360

QY 2387 AGATATGTCCTTACCAATAG---TTTACTAGTGTAGTACC-----GTCTGAGCCAG 2441
Db 361 GATATGTCCTTACCAATAGTTTACTAGTGTAGTACC-----GTCTGAGCCAG 420

QY 2442 TGATGCT---GGATTTCTGTGGAAATTTT---ATCAGAACCCAGTACTCGGGGACTGAAGG 2494
Db 421 TGATGCT---GGATTTCTGTGGAAATTTT---ATCAGAACCCAGTACTCGGGGACTGAAGG 480

QY 2495 GTGCAACCTGCAT---CAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCAAG 2552
Db 481 GTGCAACCTGCAT---CAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAAAACCAAG 540

QY 2553 TGCTAA--GCAGTGGCGGACACCATGTCCTTGAGGACAGCATGTGGAGATTGCACAGGG 2611
Db 541 TGCTAA--GCAGTGGCGGACACCATGTCCTTGAGGACAGCATGTGGAGATTGCACAGGG 600

QY 2612 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGACTCCCAATGCTATG 2671
Db 601 GCAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGACTCCCAATGCTATG 660

QY 2672 TGGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACGATGAGACCTGCCCCCTG 2731
Db 661 TGGCTCTCTCCCTTTGGCCAGTGTATGGAATGTATACGATGAGACCTGCCCCCTG 720

QY 2732 AAAATTGTTTCAAGCTACTGCTACTGTAGTCAATTTGAGGAAACAGAGCTGTGGCTGGT 2791
Db 721 AAAATTGTTTCAAGCTACTGCTACTGTAGTCAATTTGAGGAAACAGAGCTGTGGCTGGT 780

QY 2792 GTACTGATCCAGCAATACTGCAAAAGGAAATGATAGAGGTTTCCATTAAGAGACCAG 2851
Db 781 GTACTGATCCAGCAATACTGCAAAAGGAAATGATAGAGGTTTCCATTAAGAGACCAG 840

QY 2852 TGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATTTCCA 2911
Db 841 TGAAGATGCTTGGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGCTCAATTTCCA 900

QY 2912 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCACTTCACTGTCCAGCTTGCACAT 2971
Db 901 GCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCACTTCACTGTCCAGCTTGCACAT 960
```

---

```
QY 2972 GCAACGCCCAAGTAAATGCATCAATCAGAGCATCTGTGAGAAATGTGAGAACTGACCA 3031
Db 961 GCAACGCCCAAGTAAATGCATCAATCAGAGCATCTGTGAGAAATGTGAGAACTGACCA 1020

QY 3032 CAGGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAACCAATGGAGGA 3091
Db 1021 CAGGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAACCAATGGAGGA 1080

QY 3092 AATGTCAAGCCATGCAAGTGGCAACGCGCTCTCTGTGCAACACCAACACACGGGCAAGT 3151
Db 1081 AATGTCAAGCCATGCAAGTGGCAACGCGCTCTCTGTGCAACACCAACACACGGGCAAGT 1140

QY 3152 GCTTCTGCACCAACCAAGGGGCTCAAGGGGACGAGTGCAGCTATGTGAGGTGAGAAATC 3211
Db 1141 GCTTCTGCACCAACCAAGGGGCTCAAGGGGACGAGTGCAGCTATGTGAGGTGAGAAATC 1200

QY 3212 GATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTTATTGACTATCAGT 3271
Db 1201 GATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTTATTGACTATCAGT 1260

QY 3272 TCACCTTTAGTCTATCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGCTA 3331
Db 1261 TCACCTTTAGTCTATCCAGGAAGATGATCGCTATTATACAGCTATCAATTTTGTGGCTA 1320

QY 3332 CTCTGACGAAACCAACAGGGATTTGGACATGTTTCATCAATGCTGCCAAGAAATTCACAC 3391
Db 1321 CTCTGACGAAACCAACAGGGATTTGGACATGTTTCATCAATGCTGCCAAGAAATTCACAC 1380

QY 3392 TCACCATCAGCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTG 3451
Db 1381 TCACCATCAGCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGAGATGCTG 1440

QY 3452 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATGTTTCTCTAATGAGAGTTTCAAT 3511
Db 1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATGTTTCTCTAATGAGAGTTTCAAT 1500

QY 3512 TTCCAAACCAACCAATATCACTTTCTTGTGTTTATGTAGTCAATAATTCACCTGGCCCCATCA 3571
Db 1501 TTCCAAACCAACCAATATCACTTTCTTGTGTTTATGTAGTCAATAATTCACCTGGCCCCATCA 1560

QY 3572 AAATTGAGTTGCTTCTCTCAGCAAGCAATTTTATGGAACCTGGTACAGTTCTTGTGTA 3631
Db 1561 AAATTGAGTTGCTTCTCTCAGCAAGCAATTTTATGGAACCTGGTACAGTTCTTGTGTA 1620

QY 3632 CTTTCTTTCAGTTGTTTCTCTCTTGTCTGCTGGTGGCTGCTGTTTGAAGATCAAC 3691
Db 1621 CTTTCTTTCAGTTGTTTCTCTCTTGTCTGCTGGTGGCTGCTGTTTGAAGATCAAC 1680

QY 3692 AAAGTTGTTGGGCTCCAGACGTCAGAGAGCAACTTCTTCGAGAGATGCAACAGATGCCA 3751
Db 1681 AAAGTTGTTGGGCTCCAGACGTCAGAGAGCAACTTCTTCGAGAGATGCAACAGATGCCA 1740

QY 3752 GCCGTCCCTTTGCTCTGTAATGTCGCTTGGAAACAGATGAGGAGCTTCTGATCTTA 3811
Db 1741 GCCGTCCCTTTGCTCTGTAATGTCGCTTGGAAACAGATGAGGAGCTTCTGATCTTA 1800

QY 3812 TTGGGGGAGTATAAAGACTGTTCCCAACCCATTTGCACTGGAGCCGTGTTTGGCAACA 3871
Db 1801 TTGGGGGAGTATAAAGACTGTTCCCAACCCATTTGCACTGGAGCCGTGTTTGGCAACA 1860

QY 3872 AAGCGCTGCTCTCTCTGTTGTTGAGGCTCCCTCGAGGCTGGGTGGATCCCTCTC 3931
Db 1861 AAGCGCTGCTCTCTCTGTTGTTGAGGCTCCCTCGAGGCTGGGTGGATCCCTCTC 1920

QY 3932 CTGGGAGTCAAGGTTCTTGTGTCAGCCCTGGTGGACATTTCTCAGCAGATGCCGA 3991
Db 1921 CTGGGAGTCAAGGTTCTTGTGTCAGCCCTGGTGGACATTTCTCAGCAGATGCCGA 1980

QY 3992 TAGTGTACAAGGAGAGTCAAGGAGCCGTGAGAAACCGAAGACAGAGCCCTTGCACAGC 4051
Db 1981 TAGTGTACAAGGAGAGTCAAGGAGCCGTGAGAAACCGAAGACAGAGCCCTTGCACAGC 2040
```







Qy	1829	AACAAGAGAAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGCACCATGACAGAT	1888
Db	2514	ATACTAATAATTTCTTAGAGCAAAGTGCCTCTCTAAACAGCTGCTTCTGTATGACAGAT	2573
Qy	1989	GTGACCAGCACACAGATTGTATCAGCTGCGACGCCAACACCAATGACTGCGCACTGGTGC	1948
Db	2574	GTTACAGATATGCAGATTGTGCCAGCTGTACTGCTCAATACAAATGGGTGCCAATGGTGTG	2633
Qy	1949	ATGACCAATTGTGTCCCGAGGAAACACACAGCTGCTCAGAAAGGCCAGATCTCCATTTTATGGT	2008
Db	2634	ATGACAAGAAATGCAATTTTCGGCAAAATAGTAAC-----GCAGTATGCTGTCAAGAACT	2687
Qy	2009	ATGAGAAATTGCCCAAGGATAACCTATGTACTACTGTATCAAGACAGACAGCTGCAGGA	2068
Db	2688	ACACCAAAATGTC---ATGTGAGAAATCAGCAGATTTGTAAACAACTTACCAGCTGTAAAA	2744
Qy	2069	GCTGTGCCCTGGACACCAACTGCCAGTGGGAGGCCCGGAATCAGGAGTGCATTTGCCCTGC	2128
Db	2745	GCTGTTCACTAACTTGAAATTGCCAGTGGGATCAGACACAGCAGAGANTCCAGGCTTTAC	2804
Qy	2129	CCGAAAATACTGTGGCAATTTGGCTGGCAATTTGGTTGGAAACTCATGTTTGAATAATTAATA	2188
Db	2805	CAGCTCATCTTTGTGGAGAAAGATGGAGTCAATATGGGGATGCTGTGCTTTAGAGTCAATT	2864
Qy	2189	CTGCCAAGGAGAAATTATGACAACTGCTAAATTTGTTCTGTAGGAACCAACAATGCCCTTTGG	2248
Db	2865	CCAGTAGAGAAACTATGACAACTGCAAAACTTTATTGCTATATCTTTAGTGGAAATCTTG	2924
Qy	2249	CTTCTCTTACAACCCAGAAGGTAGAATTTGCTTTAAGCAGCTCGGAATTAATGCAGT	2308
Db	2925	CTTCATTTAAACAACCTCAAAGAAAGTAGAATTTGTTCT-----TGGATG	2966
Qy	2309	CATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTGGCTTCGGAAGATCAATG	2368
Db	2967	AAATACAGAAGTATACACAACAGAAAGTATCACCTGGGTGAGGCTTCGCAAGATCAATA	3026
Qy	2369	TGTCCTACTGTGTCTGGAGAGATGATGCCCAATTACAATAGTTTACTACAGTGGATGC	2428
Db	3027	TATCTTATTTGGGATGGAGACATGTCCTCTTTTACAACACAACTACAGTGGCTTC	3086
Qy	2429	CGTCTGAGCCGAGTATGCTGGAAATCTGTGGAAATTTATCAGAACCCAGTACTCGGGGAC	2488
Db	3087	CTGGCGAACCCAAATGATTTCTGGGTTTGTGCAATCTGGAAGGGCTGCAGTGGCAGGCT	3146
Qy	2489	TGAGGCTGCAAACCTGCATCAACCCATCAATGATGTGCTGTGGAAGGCTCTG-----	2542
Db	3147	TAAAGCTTAATCTTTGATCATCTATGGAATATGGCAATGGCCTGTCTGTGAAAAACCTGTTGTTA	3206
Qy	2543	---CAAACCAACAGTGTGAAGCAGTGCAGGACCAACATGTGCTTGAGCAGCAGTGTGGAG	2599
Db	3207	GTCCAAATCAAAATGCGAGGCCGTGCAAAAGCCATGCTCTCTGAGGACATCATGTTCCA	3266
Qy	2600	ATTGCACAGGGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAACGACGTGTGGACT	2659
Db	3267	ACTCTACAAGCAATGGCATGAGGTGTATGTGGTCAGCAGTACGAAACGATGTGTTGACT	3326
Qy	2660	CCAATGCCATGTGGCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGACGA	2719
Db	3327	CTAATGCCATATCATCTCTTTTCCATATGACAAATGTCTAGAGTGCCAAC---TGCCA	3383
Qy	2720	CCTGCCCCCTGAAAATTTGTTTCAGGCTACTGTAACCTGTAGTCAATTGCTTGGAGCAACAG	2779
Db	3384	CCTGCTCCCTCTAAATTTGTTCTGGATTGAAACCTCTGGACAGTGTTTGGAAACAGCCTG	3443
Qy	2780	GCTGTGCTGTGTACTGATCCACAGCAATCTGCAAGAGGAAATGCATAGAGGTTTCTT	2839
Db	3444	GATGTGGCTGTGCAATGATCCTAGTAAATACAGGAAGGACATTCATTTGAAGTTCTTT	3503
Qy	2840	ATAAAGGACCAAGTGAAGATGCTTCGCAAGGCCCTACAGGAAATTTCTATCCACAGCCCC	2899
Db	3504	CACGGGACCAATGAAG-----CTTATTGGAATGCACCAATAATGAGATGG	3548

QY	2900	TGCTCAATTCCAGCATGTGTCTAGAGACACAGATCAACTGGTCTTTTCAITTCACGTGTC	2959
Db	3549	TTCTTGACACCAATCTTTTGGCCCCAAGAAGAACTATGAGTGGTCTTTATCCAGTGTCC	3608
QY	2960	CAGCTTGCCAAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGTG	3019
Db	3609	CAGCTTGCCAGTGTAAATGGACATAGCACCTTGCAATCAATAATATGTGTGCGAACAGTGT	3668
QY	3020	AGAACCTGACACACAGGCAAGACATGCGAGACCTGTCATATCTGGCTTCTTAGGTGATCCCA	3079
Db	3669	AAAATCTCACACACAGGAAGACAGTGTCAAGATGTATGCCAGGTATATATGGAGATCCAA	3728
QY	3080	CCAATGAGGGAATATGTCAGCCATGCAAGTGCATATGGCGCACGGCTCTGTGCAACACCA	3139
Db	3729	CCAATGTGTGACAGTGCACAGCTTGTACATGCAAGTGGCCATGCAATATCTGTCAATCTGC	3788
QY	3140	ACACGGCAAGTGTCTTCGACCAACCAAGGGCGTCAAGGGGACGAGTGCACAGCTATGTG	3199
Db	3789	ACACAGAAATATGTTTCTGCACAACTAAAGGAATAAAGGTGACCAATGCCAATATATGTG	3848
QY	3200	AGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATATACATCTTCTTA	3259
Db	3849	ACTCTGAAATCGCTATGTGGTAAATCCACTTAGGGAACATGTTATACAGCCCTTTTGA	3908
QY	3260	TTGACTATCAGTTCAACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCA	3319
Db	3909	TTGATTTATCAATTTACCTTCAGCTTATTACAGGAAGATGATGGCCACCATCTGCCATAA	3968
QY	3320	ATTTTGTGGTACTCTCCGACGAACAAACAGGNTTTGGACATGTTTCATCAATGCCCTCCA	3379
Db	3969	ACTTTATAGCAACCCCAAGACAGTCGAACAAAATCTGGATATATCAATTAATGATCAAA	4028
QY	3380	AGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAG	3439
Db	4029	ACAATTTTATCTCAACATTTCTGGTGTCTGTGGTTCACACGCTGGAAACAATATCTGGG	4088
QY	3440	AAGAGATGCCGTGTGTTTCAAACCAACAATAAGGAGTCAAGAATAGTATTTCTTAATG	3499
Db	4089	AAGAGACTTCTATAGTTTCCAAAGAAATATATAAAGGAATACAGAGATAGTTTTCCTATG	4148
QY	3500	AGAGTTTGATTTTCGCAACACCCCAATATCACTTTCTTTGTTTATGTGAGTAATTTCA	3559
Db	4149	AAAAATTTAATTTAGAGAGCATCTTAACATTAATTTATGTGATGCTCAGCAACTTTT	4208
QY	3560	CCTGGCCCATCAAAATTCAGATTCGCTTCTCTCAGCACAGCAATTTTATGGACCTGTAC	3619
Db	4209	CCTGGCTATTAAATAACAGATTCGATTCATCTCACAAACAATAACAATCATGGACCTGTGC	4268
QY	3620	AGTTCTTCGTGACATTTCTCAGTTGTTTCCCTCTCTTTGCTCCCTGGTGGCTGCTGTGTTT	3679
Db	4269	AGTTTGTGTGACATTTCTCAGTTGTTTCCCTATCCCTATTTGCTGGTGGCTGCTGTGTTAT	4328
QY	3680	GGAGATCAAAACAAAGTTGTTGGGCTCCAGACGTAGAGAGCAATTTCTTCGAGAGATGC	3739
Db	4329	GGAGATCAACAACATTTGTTGGCTTCTCGACGAGAGAGCAACTGCTTCGAGAACGAC	4388
QY	3740	AACAGATGGCCAGCGCTCCCTTTGCCCTCTGTAAATGTGCGCTTGGAAACAGATGAGGAGC	3799
Db	4389	AGCAGATGGCCAGCGCTCCCTTTGCCCTCTGTGTATGTAGTCTCGAATGGAGCTCAAC	4448
QY	3800	CTCCTGATCTTTATTTGGGGAGTATAAAGACTGTGTCCAAACCCATTCGACTGGAGCGGT	3859
Db	4449	AAAACAGAGTTTCTCGAGAGGCCATTAGAGGGGCAACCCAGCCAAATTTGCCATTGAACAT	4508
QY	3860	GTTTTGGCAACAAAGCGCTGTCTCTCTGTGTGTTTGTAGAGCTCCCTCGAGGCGCTGGGTG	3919
Db	4509	GTGCTGGGAACAGAGCTGCTGTTCTGAATGTGTTTCTTTGCTATGTAGTCTCGAATGGAGCTCAG	4568
QY	3920	GCATCCCTCTCCTGCGGCAGTCAAGTCTTGCTGTGGCCAGCGCCCTGGTGGACATTTCTC	3979
Db	4569	GTGCCCCCTCCCCCTGGGCAATCAGGCCCTTGCAATTTGCAAGTGGCCCTAATAGATATTTAC	4628
QY	3980	AGCAGATGCCGATAGTGTACAGAGAGAGTCCAGAGCGCGTGTAGAAACCGGAGACGC	4036

Db 4629 AACAGAAAGCTTCAGATAGTAAAGATAAGACTTCTCGAGTCCGGAATCGAAACACC 4685

RESULT 7

US-09-893-238-8

; Sequence 8, Application US/09893238

; Patent No. US20020150973A1

; GENERAL INFORMATION:

; APPLICANT: Moore, K.

; APPLICANT: Nagle, D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY

; FILE REFERENCE: 7853-237

; CURRENT APPLICATION NUMBER: US/09/893,238

; CURRENT FILING DATE: 2001-06-27

; PRIOR APPLICATION NUMBER: 09/245,041

; PRIOR FILING DATE: 1999-02-05

; PRIOR APPLICATION NUMBER: 60/093,630

; PRIOR FILING DATE: 1998-07-21

; PRIOR APPLICATION NUMBER: 60/104,978

; PRIOR FILING DATE: 1998-10-20

; NUMBER OF SEQ ID NOS: 129

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 8

; LENGTH: 2419

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-893-238-8

Query Match 35.7%; Score 1451.8; DB 9; Length 2419;

Best Local Similarity 89.6%; Pred. No. 0;

Matches 1561; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

Qy 189 ACTAAGTGGTCTCTCTGGTTTGTGACAGATGACCTGGAATTAATAATCAAAACGNA 248

Db 677 ACTAAGTGGTCTCTCTGGTTTGTGACAGATGACCTGGAATTAATAATCAAAACGNA 736

Qy 249 GTGACGTGGTCTCAATTGAAGGACAGCAAAATAGAAATAGAAATCAATTCATTT 308

Db 737 GTGACATGGTCTCAATTGAAGGACAGCAAAATAGAAATAGAAATCAATTCATTT 796

Qy 309 TGCTACAGAGTGTAGTGGACCAATTAATAGTTAAGTGGGACTCAATTTATGACCC 368

Db 797 TGCTACAGAGTGTAGTGGACCAATTAATAGTTAAGTGGGACTCAATTTATGACCC 856

Qy 369 GCTAGTGTGCAATTTAGTGGCTCAATTTCTGACAGAGATGGCAATGACAGTGTCC 428

Db 857 TCTGATGTGCTCTTGTAGTGGCTCAATTTCTGACAGAGATGGCAATGACAGTGTCC 916

Qy 429 TGAGGTGTGTGCACATCAGGTATGCTTGTGCAATTTTGTGATGCTGCTTATATA 488

Db 917 TGAGGTGTGTGCACATCAGGTATGCTTGTGCAATTTTGTGATGCTGCTTATATA 976

Qy 489 TTTCAGTGGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 548

Db 977 TCTGATGTGCTCTTGTAGTGGCTCAATTTCTGACAGAGATGGCAATGACAGTGTCC 1036

Qy 549 AGCAGAGTGTAGATCAGTAATAGCAGCGATCTGTTGTAATGTAATGTTCTGAAACCTG 608

Db 1037 AGCAGAGTGTAGATCAGTAATAGCAGCGATCTGTTGTAATGTAATGTTCTGAAACCTG 1096

Qy 609 GAAAGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGTTTCTTCATCGAGG 668

Db 1097 GAAAGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGTTTCTTCATCGAGG 1156

Qy 669 CATCTGCAATTCAGTGTGATGTGACAGATGCTCCTGCTCTCAGACTGGCAGGCTCTGG 728

Db 1157 CATCTGCAATTCAGTGTGATGTGACAGATGCTCCTGCTCTCAGACTGGCAGGCTCTGG 1216

Qy 729 ATGTTTCAATTCCTGTGCCAGCTAACAGTCTTTTGGACTCGAGAGAAATATTTCAACTT 788

Db 1217 ATGTTTCAATTCCTGTGCCAGCTAACAGTCTTTTGGACTCGAGAGAAATATTTCAACTT 1276

Qy 789 AAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTTATGTGGGTGTGG 848

Db 1277 AAAGCTCCCGAGAGCATCTCATAAAGCTGTGGTCAATGGAAATATAATGTGGGTGTGG 1336

Qy 849 AGGATATATGTTCAACCACTCAGATTAATAACATGGTCTTAGCGTATGACCTCTCTCTAG 908

Db 1337 CGGATATATGTTCAACCACTCAGATTAATAACATGGTCTTAGCGTATGACCTCTCTAG 1396

Qy 909 GGAGTGGCTCCCACTAAACCGTCTGTGAACAATGTGGTTGTAGATATGGTCAATCTTT 968

Db 1397 GGAATGGCTCCCACTAAACCGTCTGTGAACAATGTGGTTGTAGATATGGTCAATCTTT 1456

Qy 969 GGCATTATCAAGGATAAATTTACATGTATGAGGAAATATGATTCACCTGGGATGT 1028

Db 1457 GGCATTATCAAGGATAAATTTACATGTATGAGGAAATATGATTCACCTGGGATGT 1516

Qy 1029 GACCAATGAGTGTGAGAGTCTTTTCAATTCATATGAGTCAATGGTGTGTGACCCCTAA 1088

Db 1517 GACCAATGAGTGTGAGAGTCTTTTCAATTCATATGAGTCAATGGTGTGTGACCCCTAA 1576

Qy 1089 GGCAAGGAGCAGTATGAGTGTGGGACTCTGACACATTTGTTACACTGAAGATGG 1148

Db 1577 AGCTAAGGATCAGTATGAGTGTGGGACTCTGACACATTTGTTACACTGAAGATGG 1636

Qy 1149 CCGAGTGTGATGCTGGTCACTTTTGGTCACTGCTCTCTATGGATATATAAGCAATGT 1208

Db 1637 CCGTGTGTGATGCTGGTCACTTTTGGTCACTGCTCTCTATGGATATATAAGCAATGT 1696

Qy 1209 GCAGGAATATGATTTGGATTAAGAACACATCGGAGTATATTAACACCCAGGCTGCTTGT 1268

Db 1697 GCAGGAATATGATTTGGATTAAGAACACATCGGAGTATATTAACACTCAGGCTGCTTGT 1756

Qy 1269 GCAAGGGGTTTACGGCCATAGCAGTGTTCAGCACCATTAGCAGGCGCCCTATACGTTCA 1328

Db 1757 GCAAGGGGTTTACGGCCATAGCAGTGTTCAGCACCATTAGCAGGCGCCCTATACGTTCA 1816

Qy 1329 TGTGGCTCAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTTACCGATA 1388

Db 1817 TGTGGCTCAAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTTACAGATA 1876

Qy 1389 TGATGCGATACCCAGATGTGGCCATTTCTTAAGGACAGCCGATTTTCCGTTACTTGCA 1448

Db 1877 CGATGCGATACCCAGATGTGGCCATTTCTTAAGGACAGCCGATTTTCCGTTACTTGCA 1936

Qy 1449 CACAGCTGTGATGATGAGTGGAAACCAATGCTGTGTGTTTGGGGGAAACACACAAATGACAC 1508

Db 1937 TACAGCTGTGATGATGAGTGGAAACCAATGCTGTGTGTTTGGGGGAAACACACAAATGACAC 1996

Qy 1509 ATCTATGAGCCATGCGCCAAATGCTTCTTCCAGATTTTCCGTTACTTGCA 1568

Db 1997 TTTCCATGAGCCATGCGCCAAATGCTTCTTCCGAGCTTCCGCTTATGACATGCTTG 2056

Qy 1569 TGACCGCTGTGATGATGAGTGGAAACCAATGCTGTGTGTTTGGGGGAAACACACAAATGACAC 1628

Db 2057 TGACCGCTGTGATGATGAGTGGAAACCAATGCTGTGTGTTTGGGGGAAACACACAAATGACAC 2116

Qy 1629 TTTAGAGTGTGATGATGAGTGGAAACCAATGCTGTGTGTTTGGGGGAAACACACAAATGACAC 1688

Db 2117 TTTAGAGTGTGATGATGAGTGGAAACCAATGCTGTGTGTTTGGGGGAAACACACAAATGACAC 2176

Qy 1689 CAGGAGCATCTGCTGATTTTCACTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTG 1748

Db 2177 CAGGAGCATCTGCTGATTTTCACTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTG 2236

Qy 1749 TTTAGAGTGTGATGATGAGTGGAAACCAATGCTGTGTGTTTGGGGGAAACACACAAATGACAC 1808

Db 2237 TTTAGAGTGTGATGATGAGTGGAAACCAATGCTGTGTGTTTGGGGGAAACACACAAATGACAC 2296

Qy 1809 GTGGGCTCGGACATGATGAACAAGGAAAGTTTAAATATGAAATGAGATGTTTTCCTAAAG 1868

Db 2297 CTGGGATTTGGCAATCTGAAGAACAGCAAGGAAAGTTTAAATATGAGATGTTTTCCTAAAG 2356

Qy	1869	AACCTTGACCATGACACAGATGTGACACAGCACACAGATGTTACAGCTGCGACAGCCACAC	1928
Db	2357	AACCCCTTGACCATGACAGAGTGTGACACAGCACACAGATGTTACAGCTGCGACAGCCAAATAC	2416
Qy	1929	CAA	1931
Db	2417	CAA	2419

## RESULT 8

```

US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-10

```

```
Query Match      11.4%; Score 463; DB 9; Length 1051;
Best Local Similarity 85.2%; Pred. No. 2.4e-132;
Matches 517; Conservative 0; Mismatches 90; Indels 0; Gaps 0;
```

QY	15	GGCGGCAACTGAGGCAAGCTGAGGAGGACGGCGGCGACGACGGCTCGCGGGCAG	74
Db	441	GGAGGCGAGGCCGCTCGGGTGGCGGGCGGGTGTCCGGCTCGCGCGGACGCGAGGCCAA	500
QY	75	GAGCGGCGGGCGCACCTGTGTCAACGGCGGTGCTGCAACCCCTGGCACCCGCGCAAGTCGT	134
Db	501	GGAATGTGACCGGGCCGTGTGTCAACGGCGGGCGGTGCAACCCCTGGCACCGGCCCAGTCGT	560
QY	135	CTGCCCCGCGGCTGGGTGGGCGGAGCAATGCACACCTGCGGGGGCGGCTTCAGACACTAAC	194
Db	561	CTGCCCCACGGGCTGGGTGGGCGGAGCAATGCACACCTGCGGGGGCGGCTTCAGACTAAC	620
QY	195	TGGATCTTCTGGGTTTGTGACAGATGGACCTGGGAAATTATAAATACAAAAAGAGTGCAC	254
Db	621	TGGCTCTTCTGGATTGTGTAAACAGATGGACCTGGGAAATTATAAATATAGACAGAAATGCAC	680
QY	255	GTGGCTCATTCGAAGGACAGCCAAATAGAAATATGAGACTTCGGTTTCAATCATTTTGTCTAC	314
Db	681	ATGGCTCATTGAGGACAGCCAAATAGAAATATGAGACTTCGGTTTCAACCATTTTGTCTAC	740
QY	315	AGAGTAGATTGGGACCAATTATATGTTTATGATGGGAGCTCAATTTATGCAACGCTAGT	374
Db	741	AGAAATGTAGCTGGGACCAATTATATGTTTATGATGGGAGCTCAATCTACGCAACCTCTGAT	800
QY	375	TGCTGCATTTAGTGGCCTCATTGTTCTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT	434
Db	801	TGCTGCCCTTATGTGSCCTCATTGTTCTCTGAAGAGATGCAATGAGACGGCTCCTTGAGGT	860
QY	435	TGTTGGCCACATCAGGTTATGCGTTGCTGCATTTTTTTTAGTGAATGCTGCTTATAATTTGAC	494
Db	861	CACGTGTCACTTCAGGTTATGACATGCTGCATTTTTTTCAGTGAATGCTGCTTATAATCTGAC	920
QY	495	TGGATTTAAATATTACTTTACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCGCGAGAGA	554

Db	921	TGGA	TTAATA	TACACT	TACAAATTTTGACATG	TCTCCGAA	TAAATGCTCAGGCCGAGGAGA	980
Qy	555	GTGT	TAGATCAG	TAAATAGCAG	CGATACTGTTGAATGTGAA	TGTAATGTTCTGAAAAC	TGCGAAAGG	614
Db	981	GTGT	AAGACAGTAA	CAGCAG	CGCGCTGTTAGTGTGAA	TGTAATGTTCTGAAAAC	TGCGAAAGG	1040
Qy	615	TGAAGCA	621					
Db	1041	GGCGGGA	1047					

## RESULT 9

```

US-10-198-846-5869/c
; Sequence 5869, Application US/10198846
; Publication No. US2003099974A1
;
; GENERAL INFORMATION:
;
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steilmann, Kathleen
;
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
;
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
;
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
;
; FILE REFERENCE: MRI-049
;
; CURRENT APPLICATION NUMBER: US/10/198,846
;
; CURRENT FILING DATE: 2002-07-18
;
; PRIOR APPLICATION NUMBER: 60/306,220
;
; PRIOR FILING DATE: 2001-07-18
;
; NUMBER OF SEQ ID NOS: 14084
;
; SOFTWARE: FastSeq for Windows Version 4.0
;
; SEQ ID NO 5869
;
; LENGTH: 625
;
; TYPE: DNA
;
; ORGANISM: Homo sapiens
;
; FEATURE:
;
; NAME/KEY: misc_feature
;
; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29
;
; OTHER INFORMATION: n = A,T,C or G
;
; US-10-198-846-5869

```

Query Match	6.8%;	Score	277.4;	DB	14;	Length	625;
Best Local Similarity	99.6%;	Pred. No.	9.5e-75;				
Matches	278;	Conservative	0;	Mismatches	1;	Indels	0;
						Gaps	0;
QY	2477	GTACTCGGGACACTGAAGCGCTGCACACCTGCGATCAACCCACCTCAATGGTAGTGTCTGTGAAA	2536				
Db	340	GTACTCGGGACCTGAAGCGCTGCAACCTGCACTCAACCCCACTCAATGGTAGTGTCTGTGAAA	281				
QY	2537	GGCCCTGCAAAACCCACAGTGGCTTAAGCAGTGGCGGACACCCATGTCCTTCAGACACACATGTG	2596				
Db	280	GGCCCTGCAAAACCCACAGTGGCTTAAGCAGTGGCGGACACCCATGTCCTTCAGACACACATGTG	221				
QY	2597	GAGATTGCACACAGCGGGACGCTCTTGAGTGCATGTGGTCAGCAACATCAAGACAGTGTGTGG	2656				
Db	220	GAGATTGCACACAGCGGGACGCTCTGAGTGTATGTGGTCAGCAACATCAAGACAGTGTGTGG	161				
QY	2657	ACTCCAAATGCTATGTGGCCCTCCTTCCTCTTTTGCCACAGTGTATGGAAATGGTATACGATGA	2716				
Db	160	ACTCCAAATGCTATGTGGCCCTCCTTCCTCTTTTGCCACAGTGTATGGAAATGGTATACGATGA	101				
QY	2717	GCACCTGCCCCCCCTGAAAAATGTTCAGGCTACTGTACCT	2755				
Db	100	GCACCTGCCCCCCCTGAAAAATGTTCAGGCTACTGTACCT	62				

## RESULT 10

US-09-864-761-15900  
; Sequence 15900, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharron G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.



```
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-966

Query Match      6.0%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.7e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGGC 2352
Db      |||||
246 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGGC 187

QY 2353 CTTCGGAAGATCAATGTGCTCTACTGCTGCGGAAGATATGTCGCCATTTTACAATAGT 2412
Db      |||||
186 CTTCGGAAGATCAATGTGCTCTACTGCTGCGGAAGATATGTCGCCATTTTACAATAGT 127

QY 2413 TTACTACAGTGGATGCGGCTCTGAGCCAGTGGTGGATTCCTGTGGAATTTTATCAGAA 2472
Db      |||||
126 TTACTACAGTGGATGCGGCTCTGAGCCAGTGGTGGATTCCTGTGGAATTTTATCAGAA 67

QY 2473 CCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGT 2532
Db      |||||
66 CCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2533 GAAAGG 2538
Db      |||||
6 GAAAGG 1

RESULT 14
US-10-017-754-966/c
; Sequence 966, Application US/10017754
; Publication No. US20030054363A1
; GENERAL INFORMATION:
; APPLICANT: Henderson, Robert A.
; APPLICANT: Wang, Tongtong
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Johnson, Jeffrey C.
; APPLICANT: Retter, Marc W.
; APPLICANT: Marnierakis, Margarita
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C18
; CURRENT APPLICATION NUMBER: US/10/017,754
; CURRENT FILING DATE: 2001-10-29
; NUMBER OF SEQ ID NOS: 2004
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-017-754-966

Query Match      6.0%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.7e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGGC 2352
Db      |||||
246 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGGC 187

QY 2353 CTTCGGAAGATCAATGTGCTCTACTGCTGCGGAAGATATGTCGCCATTTTACAATAGT 2412
Db      |||||
186 CTTCGGAAGATCAATGTGCTCTACTGCTGCGGAAGATATGTCGCCATTTTACAATAGT 127

QY 2413 TTACTACAGTGGATGCGGCTCTGAGCCAGTGGTGGATTCCTGTGGAATTTTATCAGAA 2472
Db      |||||
126 TTACTACAGTGGATGCGGCTCTGAGCCAGTGGTGGATTCCTGTGGAATTTTATCAGAA 67

QY 2473 CCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGT 2532
Db      |||||
66 CCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2533 GAAAGG 2538
Db      |||||
6 GAAAGG 1

RESULT 15
US-10-113-872-966/c

; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: McNabb, Andria
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C17
; CURRENT APPLICATION NUMBER: US/09/902,941
; CURRENT FILING DATE: 2001-07-10
; NUMBER OF SEQ ID NOS: 2002
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-941-966

Query Match      6.0%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.7e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGGC 2352
Db      |||||
246 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGGC 187

QY 2353 CTTCGGAAGATCAATGTGCTCTACTGCTGCGGAAGATATGTCGCCATTTTACAATAGT 2412
Db      |||||
186 CTTCGGAAGATCAATGTGCTCTACTGCTGCGGAAGATATGTCGCCATTTTACAATAGT 127

QY 2413 TTACTACAGTGGATGCGGCTCTGAGCCAGTGGTGGATTCCTGTGGAATTTTATCAGAA 2472
Db      |||||
126 TTACTACAGTGGATGCGGCTCTGAGCCAGTGGTGGATTCCTGTGGAATTTTATCAGAA 67

QY 2473 CCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGT 2532
Db      |||||
66 CCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCAATCAACCCACTCAATGGTAGTGTCTGT 7

QY 2533 GAAAGG 2538
Db      |||||
6 GAAAGG 1

RESULT 13
US-09-849-626-966/c
; Sequence 966, Application US/09849626
; Publication No. US20020197669A1
; GENERAL INFORMATION:
; APPLICANT: Bangur, Chaitanya
; APPLICANT: Fanger, Gary
; APPLICANT: Wang, Aijun
; APPLICANT: Wang, Tongtong
; APPLICANT: Switzer, Anne
; APPLICANT: McNeill, Patricia
; APPLICANT: Clapper, Jonathan
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C16
; CURRENT APPLICATION NUMBER: US/09/849,626
; CURRENT FILING DATE: 2001-05-03
; NUMBER OF SEQ ID NOS: 1926
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-849-626-966

Query Match      6.0%; Score 246; DB 9; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.7e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2293 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGGC 2352
Db      |||||
246 CTGCGAATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGGC 187
```

```

; Sequence 966, Application US/10113872
; Publication No. US20030170255A1
; GENERAL INFORMATION:
; APPLICANT: Watanabe, Yoshihiro
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Sleath, Paul R.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; FILE REFERENCE: 210121.478G19
; CURRENT APPLICATION NUMBER: US/10/113,872
; CURRENT FILING DATE: 2002-03-28
; NUMBER OF SEQ ID NOS: 2011
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-113-872-966

Query Match      6.0%; Score 246; DB 14; Length 246;
Best Local Similarity 100.0%; Pred. No. 2.7e-65;
Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2293 CTGCGAATATGCAGTCATCTCAGACATGTCCAAAGCTCACCTTAACCCCATGGGTGGGC 2352
Db 246 CTGCGAATATGCAGTCATCTCAGACATGTCCAAAGCTCACCTTAACCCCATGGGTGGGC 187

Qy 2353 CTTCCGAAGATCAATGTGTCTACTGGTGTGGGAAGATATGTCCCATTTACAAATAGT 2412
Db 186 CTTCCGAAGATCAATGTGTCTACTGGTGTGGGAAGATATGTCCCATTTACAAATAGT 127

Qy 2413 TTACTACAGTGGATGCGGTCTGAGCCAGTCATGCTGGATTCTGTGGATTTTATCAGAA 2472
Db 126 TTACTACAGTGGATGCGGTCTGAGCCAGTCATGCTGGATTCTGTGGATTTTATCAGAA 67

Qy 2473 CCCAGTACTCGGGGACTGAAGGCTGCAAGCTGCATCAACCCACTCAATGCTAGTGTCTGT 2532
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAAGCTGCATCAACCCACTCAATGCTAGTGTCTGT 7

Qy 2533 GAAAGG 2538
Db 6 GAAAGG 1

```

Search completed: March 5, 2004, 13:18:32  
Job time : 906.264 secs





Db 534 AGAGTGTAGTTGGGACCAATTATATATGTTTATGATGGGGAAGTCAAAATTATGCAACCGTAGT 593  
QY 375 TGCTGCAATTTAGTGGGCTCATTTGTTCTCAGAGAGATGGCAATGAGACTGTCCCTCAGGT 434  
Db 594 TGCTGCAATTTAGTGGGCTCATTTGTTCTCAGAGAGATGGCAATGAGACTGTCCCTCAGGT 653  
QY 435 TGTTGCCACATCAGGTTATGCCCTTGCTGCAATTTTATGATGCTGCTTAAATTTGAC 494  
Db 654 TGTTGCCACATCAGGTTATGCCCTTGCTGCAATTTTATGATGCTGCTTAAATTTGAC 713  
QY 495 TGAATTTAATATTACTTTACAGTTTTCATATGTGCCAATAAATCTGCACAGGCCGAGGAGA 554  
Db 714 TGAATTTAATATTACTTTACAGTTTTCATATGTGCCAATAAATCTGCACAGGCCGAGGAGA 773  
QY 555 GTGTAAGATCAGTAATAGCAGGATACTGTTGAAATGTGAATGTTCTGAAAACTGGAAGG 614  
Db 774 GTGTAAGATCAGTAATAGCAGGATACTGTTGAAATGTGAATGTTCTGAAAACTGGAAGG 833  
QY 615 TGAAGCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTTCTCATCGAGGCATCTG 674  
Db 834 TGAAGCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTTCTCATCGAGGCATCTG 893  
QY 675 CAATTCAGATGTGTCAGAGGATGCTCCTGCTTCTCAGACTGCGAGGGTCTCTGGATGTC 734  
Db 894 CAATTCAGATGTGTCAGAGGATGCTCCTGCTTCTCAGACTGCGAGGGTCTCTGGATGTC 953  
QY 735 AGTTCCTGTACAGCTTAAACAGTCAATTTGGACTCGAGAGGATATTTCTAACTTAAAGCT 794  
Db 954 AGTTCCTGTACAGCTTAAACAGTCAATTTGGACTCGAGAGGATATTTCTAACTTAAAGCT 1013  
QY 795 CCCCAGAGATCTCAATAAGCTGTGTCAATGGAACATTTATGGGTTGTTGGAGGATA 854  
Db 1014 CCCCAGAGATCTCAATAAGCTGTGTCAATGGAACATTTATGGGTTGTTGGAGGATA 1073  
QY 855 TATGTTCAACCACTCAGATTAATCAATGTTCTAGCGTATGACCTTCTTAGGGAGTG 914  
Db 1074 TATGTTCAACCACTCAGATTAATCAATGTTCTAGCGTATGACCTTCTTAGGGAGTG 1133  
QY 915 GCTTCCACTAAACCGTTCTGTGAACAACTGTGTTGTAGATATGTTCACTTCTTGCCATT 974  
Db 1134 GCTTCCACTAAACCGTTCTGTGAACAACTGTGTTGTAGATATGTTCACTTCTTGCCATT 1193  
QY 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTCAACTGGGAATGTGACCAA 1034  
Db 1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTCAACTGGGAATGTGACCAA 1253  
QY 1035 TGAATTTGAGATTTTTCATTAATCAATGAGTCACTGGGTTGTTGACCCCTAAGGCCAAA 1094  
Db 1254 TGAATTTGAGATTTTTCATTAATCAATGAGTCACTGGGTTGTTGACCCCTAAGGCCAAA 1313  
QY 1095 GGAGCAGTATGCAATGGTTGGGCACTCTGCACACATTTGTTACACTGAAGAAATGGCCGAGT 1154  
Db 1314 GGAGCAGTATGCAATGGTTGGGCACTCTGCACACATTTGTTACACTGAAGAAATGGCCGAGT 1373  
QY 1155 GGTCACTGCTGTCATCTTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGACAGA 1214  
Db 1374 GGTCACTGCTGTCATCTTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGACAGA 1433  
QY 1215 ATATGATTTGGATAAGAAACATGAGTATATTAACACACCCAGGGTGCCTTTGTGCAAGG 1274  
Db 1434 ATATGATTTGGATAAGAAACATGAGTATATTAACACACCCAGGGTGCCTTTGTGCAAGG 1493  
QY 1275 GGTTACGGCCATPAGCAGTGTTTAACAATAGGACCGAGGCCCTATAGCTTTCATGGTGG 1334  
Db 1494 GGTTACGGCCATPAGCAGTGTTTAACAATAGGACCGAGGCCCTATAGCTTTCATGGTGG 1553  
QY 1335 CTACAAGGCTTTTCAGTGCCAAATAGTACCGGTTGACAGATGATCTCTACCGATATGATGT 1394  
Db 1554 CTACAAGGCTTTTCAGTGCCAAATAGTACCGGTTGACAGATGATCTCTACCGATATGATGT 1613  
QY 1395 GGATACCCAGATGTGGACCAATTTCTTAAGGACAGCCGATTTTTCCGTACTCTGCAACAGC 1454  
Db 1614 GGATACCCAGATGTGGACCAATTTCTTAAGGACAGCCGATTTTTCCGTACTCTGCAACAGC 1673

QY 1455 TGTGATAGTCAGTGGAAACCAATGCTGTGTTTGGGGGAAACACACAAATGACACATCTAT 1514  
Db 1674 TGTGATAGTCAGTGGAAACCAATGCTGTGTTTGGGGGAAACACACAAATGACACATCTAT 1733  
QY 1515 GAGCCATGGCGCCAAATGCTTCTTTCAGATTTTCATGGCCCTATGACATTGGCTGTGACCG 1574  
Db 1734 GAGCCATGGCGCCAAATGCTTCTTTCAGATTTTCATGGCCCTATGACATTGGCTGTGACCG 1793  
QY 1575 CTGTCAGTGTCTTCCCAAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1634  
Db 1794 CTGTCAGTGTCTTCCCAAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGC 1853  
QY 1635 AGTCTTACAAACAGCACCATGTATGTTCGGTGGTTTCAATAGTCTCTCCTCAGCGA 1694  
Db 1854 AGTCTTACAAACAGCACCATGTATGTTCGGTGGTTTCAATAGTCTCTCCTCAGCGA 1913  
QY 1695 CATCTCTGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTAGC 1754  
Db 1914 CATCTCTGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTTAGC 1973  
QY 1755 AGCAGGACCTGCTGTTTCGGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCTGGGC 1814  
Db 1974 AGCAGGACCTGCTGTTTCGGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCTGGGC 2033  
QY 1815 GCTGGCAACTGTATGAAACAAGAAAGTTAAATCAGATGTTTTCAAAAAGAACTCT 1874  
Db 2034 GCTGGCAACTGTATGAAACAAGAAAGTTAAATCAGATGTTTTCAAAAAGAACTCT 2093  
QY 1875 TGACCATGACAGATGTGACACAGCACACAGATGTTACAGCTGCAACAGCCAAACCAATGA 1934  
Db 2094 TGACCATGACAGATGTGACACAGCACACAGATGTTACAGCTGCAACAGCCAAACCAATGA 2153  
QY 1935 CTGCCACTGTGTCAATGACCAATTTGTCGCCAGAAACCAAGCTGCTCAGAGAGCCAGAT 1994  
Db 2154 CTGCCACTGTGTCAATGACCAATTTGTCGCCAGAAACCAAGCTGCTCAGAGAGCCAGAT 2213  
QY 1995 CTCCTATTTTAGTGTATGAGAAATTCGCCCAAGGATAACCTTATCTACTGTAAACAGAA 2054  
Db 2214 CTCCTATTTTAGTGTATGAGAAATTCGCCCAAGGATAACCTTATCTACTGTAAACAGAA 2273  
QY 2055 GACAGCTGACAGAGCTGTCCCTGACACAGAACTGCCAGTGGAGGCCCGGATCAGGA 2114  
Db 2274 GACAGCTGACAGAGCTGTCCCTGACACAGAACTGCCAGTGGAGGCCCGGATCAGGA 2333  
QY 2115 GTGCATTTGCCCTGCCGAAAATATCTGTGCAATTTGGTGGCAATTTGGTTGGAACTCATG 2174  
Db 2334 GTGCATTTGCCCTGCCGAAAATATCTGTGCAATTTGGTGGCAATTTGGTTGGAACTCATG 2393  
QY 2175 TTTGAAAATTTACTCTGCCAAGGAGAAATTTATGCAATTTGTTCTGTAGGAAACCA 2234  
Db 2394 TTTGAAAATTTACTCTGCCAAGGAGAAATTTATGCAATTTGTTCTGTAGGAAACCA 2453  
QY 2235 CAATGCCCCCTTTTGGCTTCTCTTACACCCAGAAAGGTAGAATTTGCTTAAAGCAGCT 2294  
Db 2454 CAATGCCCCCTTTTGGCTTCTCTTACACCCAGAAAGGTAGAATTTGCTTAAAGCAGCT 2513  
QY 2295 GCGAATTAATCAAGTCACTCAGAGCATGTCCAAAGCTCACCTTAAACCCATGGGTGGCCT 2354  
Db 2514 GCGAATTAATCAAGTCACTCAGAGCATGTCCAAAGCTCACCTTAAACCCATGGGTGGCCT 2573  
QY 2355 TCGGAAGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCTCAATTTACAAATAGTTT 2414  
Db 2574 TCGGAAGATCAATGTGTCTTACTGTGCTGGGAAGATATGTCCCTCAATTTACAAATAGTTT 2633  
QY 2415 ACTACAGTGGATGCGCTCTGAGCCCAAGTGTGATGCTGTGAATTTCTGTGAATTTTATCAGAAC 2474  
Db 2634 ACTACAGTGGATGCGCTCTGAGCCCAAGTGTGATGCTGTGAATTTCTGTGAATTTTATCAGAAC 2693  
QY 2475 CAGTACTCGGGGATGAGGCTGCACTCAACCCATCACTCAATGGTAGTCTGTGTGA 2534  
Db 2694 CAGTACTCGGGGATGAGGCTGCACTCAACCCATCACTCAATGGTAGTCTGTGTGA 2753

Qy	2535	AAGCCTGCAAAACCA	CAGTGCCTAAG	CAGTGC	CGGACACCA	TGTCCTTTGAGGACAGCATG	2594
Db	2754	AAGCCTGCAAAACCA	CAGTGCCTAAG	CAGTGC	CGGACACCA	TGTCCTTTGAGGACAGCATG	2813
Qy	2595	TGGAGATTGCAC	CAGCGGACGCTC	TGAGTGCAT	TGTCGAGCAA	CATGAAGCAGTGTGT	2654
Db	2814	TGGAGATTGCAC	CGGACGCTCT	TGAGTGCAT	TGTCGAGCAA	CATGAAGCAGTGTGT	2873
Qy	2655	GGACTCCAAATG	CTATGTGSCCTCT	TTCCTTTTGGCCAGT	GTATGGAATGGTATACGAT	2714	
Db	2874	GGACTCCAAATG	CTATGTGSCCTCT	TTCCTTTTGGCCAGT	GTATGGAATGGTATACGAT	2933	
Qy	2715	GAGCACCTGCC	CCCTCAAATTTG	TCAGGCTACT	GTFAACTGTAGT	CATTGCTTGAGCA	2774
Db	2934	GAGCACCTGCC	CCCTCAAATTTG	TCAGGCTACT	GTFAACTGTAGT	CATTGCTTGAGCA	2993
Qy	2775	ACCAGGCTGTG	GGCTGTACTGAT	CCAGCAATPAT	TGGCAAGGGAAATGCATAGAGG	2834	
Db	2994	ACCAGGCTGTG	GGCTGTACTGAT	CCAGCAATPAT	TGGCAAGGGAAATGCATAGAGG	3053	
Qy	2835	TTCTATAAAGG	ACAGTGAAGTGC	CTTCGACGCCCT	TACAGGAAATTTCTATCCACA	2894	
Db	3054	TTCTATAAAGG	ACAGTGAAGTGC	CTTCGACGCCCT	TACAGGAAATTTCTATCCACA	3113	
Qy	2895	GCCCTGCTCAA	TTCCAGCATGTCT	TAGAGGACAGCAGAT	TACAACTGGTCTTTTCAATCA	2954	
Db	3114	GCCCTGCTCAA	TTCCAGCATGTCT	TAGAGGACAGCAGAT	TACAACTGGTCTTTTCAATCA	3173	
Qy	2955	CTGTCCAGCTT	GCCAAATGCAAC	GGCCACAGTAAT	TGCAATCAGACGATCTGTGAGAA	3014	
Db	3174	CTGTCCAGCTT	GCCAAATGCAAC	GGCCACAGTAAT	TGCAATCAGACGATCTGTGAGAA	3233	
Qy	3015	GTGTGAAAC	CTGACCA	CAGGCAAGCACT	CGAGACCTGCATATCTGGCTTCTACGGTGA	3074	
Db	3234	GTGTGAAAC	CCTGACCA	CAGGCAAGCACT	CGAGACCTGCATATCTGGCTTCTACGGTGA	3293	
Qy	3075	TCCCAACCAAT	TGGAGGAAATGT	CAGCCATGCAAT	TGGGCAACGCGTCTCTGTGCAA	3134	
Db	3294	TCCCAACCAAT	TGGAGGAAATGT	CAGCCATGCAAT	TGGGCAACGCGTCTCTGTGCAA	3353	
Qy	3135	CACCAACACGGG	CAAGTGCTT	CTGCACACCA	CAGGGCGTCAAGGGGACAGATGCCAGCT	3194	
Db	3354	CACCAACACGGG	CAAGTGCTT	CTGCACACCA	CAGGGCGTCAAGGGGACAGATGCCAGCT	3413	
Qy	3195	ATGTGAGGTAGA	AAATCGATACCA	AGAAACCCCTCT	CAGAGGAACATGTTTATTATCTCT	3254	
Db	3414	ATGTGAGGTAGA	AAATCGATACCA	AGAAACCCCTCT	CAGAGGAACATGTTTATTATCTCT	3473	
Qy	3255	TCCTATTGACT	ATCAGTTCACCTT	TAGTCTATCC	CAGGAGATGATCGCTATTATACACG	3314	
Db	3474	TCCTATTGACT	ATCAGTTCACCTT	TAGTCTATCC	CAGGAGATGATCGCTATTATACACG	3533	
Qy	3315	TATCAAATTTTG	GTGCTACTCTCTG	ACCAACAAACAGGGAT	TTTGACATGTTTCATCAATGC	3374	
Db	3534	TATCAAATTTTG	GTGCTACTCTCTG	ACCAACAAACAGGGAT	TTTGACATGTTTCATCAATGC	3593	
Qy	3375	CTCCAAGAAAT	TTCAACCTCAACAT	CTCGGCTGCCAGT	TTCTCAGCTCGAACCCAGGC	3434	
Db	3594	CTCCAAGAAAT	TTCAACCTCAACAT	CTCGGCTGCCAGT	TTCTCAGCTCGAACCCAGGC	3653	
Qy	3435	TGGAGAAGAGAT	GCCTGTGTTT	TCAAAAACCAACAT	TAAAGAGTACAAAGATAGTTTCTC	3494	
Db	3654	TGGAGAAGAGAT	GCCTGTGTTT	TCAAAAACCAACAT	TAAAGAGTACAAAGATAGTTTCTC	3713	
Qy	3495	TAATGAGAAGTT	GTGATTTTGC	AAACCAACCAATATCA	CTTTCTTTGTTTATGTCAGTAA	3554	
Db	3714	TAATGAGAAGTT	GTGATTTTGC	AAACCAACCAATATCA	CTTTCTTTGTTTATGTCAGTAA	3773	
Qy	3555	TTTTCACCTG	GGCCCATCAAAAT	TCAGATGTCCT	CTCTCAGCACAGCAATTTTATGACCT	3614	
Db	3774	TTTTCACCTG	GGCCCATCAAAAT	TCAGATGTCCT	CTCTCAGCACAGCAATTTTATGACCT	3833	
Qy	3615	GGTACAGTCT	TCGTGACTTTCT	TTTCAGTGTGTTT	CTCTTTGCTCTGGTGCTGCTGT	3674	

```

Db      3834  GGTCACAGTCTCTCGTGACTTCTTCAGTGTCTTCCTCTCTTGTCTCTGFGGCTGCTGT 3893
Qy      3675  GGTTTGGAAGATCAAAACAAAAGTTGTTGGGCGCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3734
Db      3894  GGTTTGGAAGATCAAAACAAAAGTTGTTGGGCGCTCCAGACGTAGAGAGCAACTTCTTCGAGA 3953
Qy      3735  GATGCAACAGATGGCGCAGCGCTCCCTTTGCGCTCTGTAAATGTCCCTTGGAAAAAGATGA 3794
Db      3954  GATGCAACAGATGGCGCAGCGCTCCCTTTGCGCTCTGTAAATGTCCCTTGGAAAAAGATGA 4013
Qy      3795  GGAGCCTCCCTGATCTTATTTGGGGGAGTATAAAGACATGTTCCCAAAACCCATTGCACATGGA 3856
Db      4014  GGAGCCTCCCTGATCTTATTTGGGGGAGTATAAAGACATGTTCCCAAAACCCATTGCACATGGA 4073
Qy      3855  GCGGTGTTTTGGCAACAAACCGCTGCTCTCTGTGTTGTGAGGTCCTCTCGAGGCGCT 3914
Db      4074  GCGGTGTTTTGGCAACAAACCGCTGCTCTCTGTGTTGTGAGGTCCTCTCGAGGCGCT 4133
Qy      3915  GGGTGGCATCCCTCCTCCTCGGCGAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 3974
Db      4134  GGGTGGCATCCCTCCTCCTCGGCGAGTCAGGTCTTGCTGTGGCCAGCGCCCTGGTGGACAT 4193
Qy      3975  TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCCTGAGAAACCGGAAGCA 4034
Db      4194  TTCTCAGCAGATGCCGATAGTGTACAAGGAGAAGTCAGGAGCCCTGAGAAACCGGAAGCA 4253
Qy      4035  GCAGCCCCCTGCACAGCGCTGGACCTGCATCTGA 4068
Db      4254  GCAGCCCCCTGCACAGCGCTGGACCTGCATCTGA 4287

RESULT 2
US-09-245-041-16
; Sequence 16, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagie, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-16

```

	Query Match	86.0%;	Score 3499.8;	DB 3;	Length 4072;
	Best Local Similarity	98.8%;	Pred. No. 0;		
	Matches 3525;	Conservative 0;	Mismatches 42;	Indels 0;	Gaps 0
QY	15	GGCGGCAACTGAGGCAAGGCTGAGGAGGAGGACGGCGCGACGGCAGCGCTCGCGGGCAG	74		
Db	234	GGAGCCGAGGCCGCTGCGGTGGCGGGCGGGTGTCCGGCTCGCGCGCAGCCGAGGCCAA	293		
QY	75	GAGCGGGGGCCGCACTGTGTCAACGGCGGTGCTGTGCACCCCTGGCAACCGGCCAGTGGCT	134		
Db	294	GGAATGTCAACGGGCGGTGTCAACGGCGGTGCTGTGCACCCCTGGCAACCGGCCAGTGGCT	353		
QY	135	CTGCGCCGCGCGCTGGGTGGCGGAGCAATGCACACACTGCGGGGGCCGCTTCAGACTAAC	194		
Db	354	CTGCGCCGCGCGCTGGGTGGCGGAGCAATGCACACACTGCGGGGGCCGCTTCAGACTAAC	413		
QY	195	TGGATCTTCTGGGTTTGTGCAGAGATGACCTCGGAAATTTATAAATACAAAGCGAATGGCAC	254		



QY 2415 ACTACAGTGGATGCCGCTCTGAGCCAGTGAATCTGTGAAATTTTATCAGAAC 2474  
DB 2634 ACTACAGTGGATGCCGCTCTGAGCCAGTGAATCTGTGAAATTTTATCAGAAC 2693  
QY 2475 CAGTACTGGGAGCTGAAGGCTGGAACCTGCAATCAACCCACTCAATGTGTCTGTGA 2534  
DB 2694 CAGTACTGGGAGCTGAAGGCTGGAACCTGCAATCAACCCACTCAATGTGTCTGTGA 2753  
QY 2535 AAGGCTGCAACACCAAGTCTAAGCAGTGCAGGACACCATGTGCTTGGAGCAGCATG 2594  
DB 2754 AAGGCTGCAACACCAAGTCTAAGCAGTGCAGGACACCATGTGCTTGGAGCAGCATG 2813  
QY 2595 TGGAGATGCAACCAAGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2654  
DB 2814 TGGAGATGCAACCAAGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGT 2873  
QY 2655 GCATCCCAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGAT 2714  
DB 2874 GCATCCCAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACGAT 2933  
QY 2715 GAGCACCTGCCCCCTGAAAAATGTTTCAGGCTACTGTACTCTGATCTTGGAGCA 2774  
DB 2934 GAGCACCTGCCCCCTGAAAAATGTTTCAGGCTACTGTACTCTGATCTTGGAGCA 2993  
QY 2775 ACCAGCTGTGGCTGGTGTACTGATCCAGCAATACTGCGAAGGAAATGCATAGAGG 2834  
DB 2994 ACCAGCTGTGGCTGGTGTACTGATCCAGCAATACTGCGAAGGAAATGCATAGAGG 3053  
QY 2835 TTCTTATAAGGACCAAGTGAAGTGCCTTGGCAAGCCCTACAGGAAATTTCTATCCACA 2894  
DB 3054 TTCTTATAAGGACCAAGTGAAGTGCCTTGGCAAGCCCTACAGGAAATTTCTATCCACA 3113  
QY 2895 GCCCTGTCTCAATTCAGCATGTGCTAGAGACAGCAGATACAACTGGTCTTCAATCA 2954  
DB 3114 GCCCTGTCTCAATTCAGCATGTGCTAGAGACAGCAGATACAACTGGTCTTCAATCA 3173  
QY 2955 CTGTCCAGCTGCCAATGCAAGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAA 3014  
DB 3174 CTGTCCAGCTGCCAATGCAAGCCACAGTAAATGCAATCAATCAGACATCTGTGAGAA 3233  
QY 3015 GTGTGAGAACCTGACCAAGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGA 3074  
DB 3234 GTGTGAGAACCTGACCAAGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGA 3293  
QY 3075 TCCCAACCAATGAGGGAATGTCAGCATGCAATGCAATGCAATGCAATGCAATGCAATG 3134  
DB 3294 TCCCAACCAATGAGGGAATGTCAGCATGCAATGCAATGCAATGCAATGCAATGCAATG 3353  
QY 3135 CACCAACACGGGCAAGTCTTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCAGCT 3194  
DB 3354 CACCAACACGGGCAAGTCTTCTGCAACCAAGGCGCTCAAGGGGACGAGTGCAGCT 3413  
QY 3195 ATGTGAGGTAGAAAATCGAATACCAAGGAAACCTCTCAGAGGAAACATGTTATATATCTCT 3254  
DB 3414 ATGTGAGGTAGAAAATCGAATACCAAGGAAACCTCTCAGAGGAAACATGTTATATATCTCT 3473  
QY 3255 TCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTAACACAGC 3314  
DB 3474 TCTTATGACTATCAGTTCACCTTTAGTCTATCCAGGAGATGATCGCTATTAACACAGC 3533  
QY 3315 TATCAATTTTGTGGCTACTCTGAGCAACAAACAGGATTTGGACATGTTTCATCAATGC 3374  
DB 3534 TATCAATTTTGTGGCTACTCTGAGCAACAAACAGGATTTGGACATGTTTCATCAATGC 3593  
QY 3375 CTCGAGATTTTCAACCTCAATCATCTGGGCTCCAGTCTTCTCAGCTGGAAACCCAGGC 3434  
DB 3594 CTCGAGATTTTCAACCTCAATCATCTGGGCTCCAGTCTTCTCAGCTGGAAACCCAGGC 3653  
QY 3435 TGGAGAGAGATGCCCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3494  
DB 3654 TGGAGAGAGATGCCCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTC 3713

QY 3495 TAATGAGAGTTCGATTTGCGAACCAACCAATATACATTTCTTTTATGTCAGTAA 3554  
DB 3714 TAATGAGAGTTCGATTTGCGAACCAACCAATATACATTTCTTTTATGTCAGTAA 3773  
QY 3555 TTTCACCTGGCCCATCAAAATTCAGAT 3581  
DB 3774 TTTCACCTGGCCCATCAAAATTCAGAT 3800

RESULT 3  
US-09-245-041-1  
; Sequence 1, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; EARLIER FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8827  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-1

Query Match 81.8%; Score 3326.4; DB 3; Length 8827;  
Best Local Similarity 91.1%; Pred. No. 0;  
Matches 3534; Conservative 0; Mismatches 346; Indels 0; Gaps 0;

QY 189 ACTAAGTGGATCTCTGGGTTTGTGACAGATGACCTGGAAATATATAATACAAACGAA 248  
DB 677 ACTAAGTGGATCTCTGGGTTTGTGACAGATGACCTGGAAATATATAATACAAACGAA 736  
QY 249 GTGACGTGGCTCATTTGAAGCAGCAGCAATAGAAATAGAGACTTGTGTTCAATCATTT 308  
DB 737 GTGACGTGGCTCATTTGAAGCAGCAGCAATAGAAATAGAGACTTGTGTTCAATCATTT 796  
QY 309 TGCTACAGGTGTAGTGGGACCATTTATATGTTATGATGGGACTCAATTTATGACC 368  
DB 797 TGCTACAGGTGTAGTGGGACCATTTATATGTTATGATGGGACTCAATTTATGACC 856  
QY 369 GCTAGTTGCTGCAATTTAGTGGGCTCATTTGTTCTGAGAGATGGCAATGAGACTGTGCC 428  
DB 857 TCTGATTTGCTGCTTTAGTGGGCTCATTTGTTCTGAGAGATGGCAATGAGACTGTGCC 916  
QY 429 TGAGGTTGTTGCCACATCAGGTTATGCTTGTGCTGCAATTTTATGATGATGCTGCTTATAA 488  
DB 917 TGAGGTTGTTGCCACATCAGGTTATGCTTGTGCTGCAATTTTATGATGATGCTGCTTATAA 976  
QY 489 TTTGACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 548  
DB 977 TTTGACTGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1036  
QY 549 AGGAGGTGTAAGATCAGTAAATAGCAGCAGTACTGTTGAAATGTAATGTTTCTGAAACTG 608  
DB 1037 AGGAGGTGTAAGATCAGTAAATAGCAGCAGTACTGTTGAAATGTAATGTTTCTGAAACTG 1096  
QY 609 GAGAGGTGAGACATGTCACATCTCTCAGTACAGCACTGTTGTTTCTCCTCATCAGG 668  
DB 1097 GAGAGGTGAGACATGTCACATCTCTCAGTACAGCACTGTTGTTTCTCCTCATCAGG 1156  
QY 669 CATCTGCAATTCAGTGTATGATCTCAGAGGATCTCTGCTTCTCAGACTGGCAGGTTCTG 728  
DB 1157 CATCTGCAATTCAGTGTATGATCTCAGAGGATCTCTGCTTCTCAGACTGGCAGGTTCTG 1216

QY 729 ATGTTCAAGTTCCTGTACAGCTAAACAGTCAATTTGGACTCGAGGAGAAATATCTAACTT 788  
Db 1217 ATGTTCAAATCTCTGTGCCAGCTAAACAGTCTTTTGGACTCGAGGAGAAATATCTCAATTT 1276  
QY 789 AAGCTCCCGAGAGCATCTCAATAAGCTGTGTCATGGAACATATGTTGGGTTGTTGG 848  
Db 1277 AAGCTTCCCGAGAGCTCTCAATAAGCTGTGTCATGGAACATATGTTGGGTTGTTGG 1336  
QY 849 AGGATATATGTTCAACCACTCAGATATATACATGTTCTTAGCGTATGACCTTGTCTTAG 908  
Db 1337 CGGATATATGTTCAACCACTCAGATATATACATGTTCTTAGCGTATGACCTTGTCTTAG 1396  
QY 909 GGAGTGGCTTCACCTAAACCGTTCTGTGAACATGTTGTTAGATATGTTCACTCTTT 968  
Db 1397 GGAATGGCTTCACCTAAACCACTCTGTGAACATGTTGTTAGATATGTTCACTCTTT 1456  
QY 969 GGCATTATACAGGATATAAATTTACATGTTAGGAGGAAATTTGATTTCAACAGGGAATGT 1028  
Db 1457 GGCATTATACAGGATATAAATTTACATGTTAGGAGGAAATTTGATTTCAACAGGGAATGT 1516  
QY 1029 GACCAATAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGGTGTTGTTGACCCCTAA 1088  
Db 1517 GACCAATAGTTGAGAGTTTTCATATTCATTAATGAGTCATGGGTGTTGTTGACCCCTAA 1576  
QY 1089 GGCAGGAGGAGCATATGAGTGTGTTGGGACCTCTGCACATGTTTACACTGAAGAAATGG 1148  
Db 1577 AGCTAAGGATCAGTATGAGTGTGTTGGACCTCTGCACATGTTTACACTGGCATCTGG 1636  
QY 1149 CCGAGTGGTCATGCTGGTCATCTTTGGTCATCTGCCCTCTCTATGATATATAAGCAATGT 1208  
Db 1637 CCGTGTGGTCATGCTGGTCATCTCGGTCAATTCGCCCTCTATGATATATAAGCGTTGT 1696  
QY 1209 GCAGGAATATGTTGGATAAGAACACATGAGTATATTACACACCGGGTGCCCTTGT 1268  
Db 1697 GCAGGAATATGACTTGGAAAGAACACATGAGTATATTACACTCTCAGGGTGCTCTTGT 1756  
QY 1269 GCAAGGGGTTACGGCCATAGCAGTGTTTACGACCATAGGACAGGGCCCTATAGTTTCA 1328  
Db 1757 GCAAGGGGTTATGGCCACAGTAGTGTTTATGATGACAGGACCAAGGCTCTGTACGTTCA 1816  
QY 1329 TGGTGGCTACAGGCTTTTCAGTGCCMAATAGTACCGGCTTGAGATGATCTCTACCGATA 1388  
Db 1817 TGGTGGCTACAGGCTTTTCAGTGCCMAATAGTACCGGCTTGAGATGATCTCTACCGATA 1876  
QY 1389 TGAATGGGATACCCAGATGGAACCACTCTTAAGGACAGCCGATTTTCCGTTACTTGCA 1448  
Db 1877 CGATGTGGATATCTCAGATGTGACCACTCTTAAGGACAGCCGATTTTCCGTTACTTGCA 1936  
QY 1449 CACAGTGTGATGAGTGGAAACCATGCTGTGTTGGGGGAAACACACAAATGACAC 1508  
Db 1937 TACAGCTGTGATGAGTGGAAACCATGCTGTGTTGGGGGAAACACACAAATGACAC 1996  
QY 1509 ATCTAGGCCATGGGCGCAATGCTTCTTCAGATTTTCAGGCTATGACATTTGCTGCTG 1568  
Db 1997 TTCATGAGCCACGGTGCMAATGCTTCTTCGAGCTTCATGGCTTATGACATTTGCTGCTG 2056  
QY 1569 TGAACCGCTGTGAGTGTTCAGGACCTGATCTCCACCATGATGTCAACAGATTTGGGCA 1628  
Db 2057 TGAACCGATGTGAGTGTTCAGGACCTGATCTCCACCATGATGTCAACAGATTTGGGCA 2116  
QY 1629 TTCAGCAGCTTACAAACAGCAGCATGATGTTGTTGCGGCTTCAATAGTCTCTCTCT 1688  
Db 2117 TTCAGCAGCTTGTGTAACAGCAGCATGATGTTGTTGCGGCTTCAACAGCCTCTCTCTCT 2176  
QY 1689 CAGCAGATCTCTGTTATTCACCTCGGAACAGTGTGATGCGCATGCGAGTGAAGCGCTTG 1748  
Db 2177 CAGTGAAGCTTGTGTTTACCTCGGAGCAGTGGGATGACACCGCAGTGAAGCTGCTG 2236  
QY 1749 TTTAGCAGCAGGACCTGTTATTCGGTGTGTTGGAAACAGGCTGCTCAGTGTATCTC 1808  
Db 2237 TGTGGCAGCAGGACCTGTTATTCGGTGTGTTGGAAACAGGCTGCTCAGTGTATCTC 2296  
QY 1809 GTGGGCGCTGGCACTGATGAAACAGAGAAAGATTTAAATCAGAATGTTTTCCTCAAAAG 1868

Db 2297 CTGGAGTGTGGCACTGAAGAACAGAGAAAGTTAAATACAGATGTTTTTCTTAAAG 2356  
QY 1869 AACTCTTGACCATGACAGATGTGACAGACACAGATTTGTACAGCTGACAGCCCAACAC 1928  
Db 2357 AACCTTGACCATGACAGATGTGACAGACACAGATTTGTACAGCTGACAGCCCAATAC 2416  
QY 1929 CAATGACTGACCTGTGCAATGCAATGTGTGCCAGGAACACAGCTGCTCAGAAAG 1988  
Db 2417 CAATGACTGACCTGTGCAATGCAATGTGTGCCAGGAACACAGCTGCTCAGAAAG 2476  
QY 1989 CCAGATCTCATTTTAGGTATGAGAAATGCCCAAGGATAACCTATGTACTACTGTAA 2048  
Db 2477 CCAGATCTCATTTAGGTATGAGAAATGCCCAAGGATAACCTATGTACTACTGTAA 2536  
QY 2049 CAAGAAGCAGCTGACAGGAGCTGTGCCCTGGACACAGAACTGCGAGTGGAGCCCGGAA 2108  
Db 2537 TAAGAAACCCAGCTGACAGGAGCTGTGCCCTAGACACAGAACTGCGAGTGGAGCCCGGAA 2596  
QY 2109 TCAGGAGTCAATTTGCCCTGCCGAAATATCTGTGCATTTGGCTGCAATTTGTTTGGAAA 2168  
Db 2597 TCAGGAGTCAATTTGCCCTGCCGAAATATCTGTGCATTTGGCTGCAATTTGTTTGGAAA 2656  
QY 2169 CTCATGTTTGAATAATTTACTACTGCCAAGGAGAAATATGACAATGCTTAAATTTGTTGTAG 2228  
Db 2657 CTCGTTGTTGAATAATTTACTACTGCCAAGGAGAAATATGACAATGCTTAAATTTGTTGTAG 2716  
QY 2229 GAACCAATATGCTTTTGGCTTCTCTTACACCCAGAGAGGATAGATTTGTTTCTTAA 2288  
Db 2717 GAACCAATATGCTTTTGGCTTCTCTTACACCCAGAGAGGATAGATTTGTTTCTTAA 2776  
QY 2289 GCAGCTCGCAATATGCAATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCCATGGGT 2348  
Db 2777 GCAGCTCGCAATATGCAATCTCAGAGCATGTCCAAAGCTCACCTTAAACCCCATGGGT 2836  
QY 2349 CGGCTTTGGGAAAGATCAATGTGCTTCTTACAGTGTGCTGGAGAGATATGTCCTTACAA 2408  
Db 2837 TGGTCTTTGGGAAAGATCAATGTGCTTCTTACAGTGTGCTGGAGAGATATGTCCTTACAA 2896  
QY 2409 TAGTTTACTACAGTGTGCTGAGCCAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTG 2468  
Db 2897 TAGTTTACTACAGTGTGCTGAGCCAGTGTGCTGAGTGTGCTGAGTGTGCTGAGTGTGCTG 2956  
QY 2469 AGAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCATCAACCCACTCAATGGTGTGT 2528  
Db 2957 AGAGCTTAGTACTCGGGGATTAAGGCTGCAACCTGCATCAACCCCTCTCAATGGCAGCT 3016  
QY 2529 CTGTGAAGGCTTGCACACACAGTGTGAGCAGTGTGAGCAGCAGTGTGCTTGAAGGAC 2588  
Db 3017 CTGTGAAGGCTTGCACACACAGTGTGAGCAGTGTGAGCAGCAGTGTGCTTGAAGGAC 3076  
QY 2589 AGCATGTGAGATTTGACACAGCGGAGCTGTGAGTGTGCTGAGTGTGCTGAGCAGCAGTGTG 2648  
Db 3077 AGCATGTGAGATTTGACACAGCGGAGCTGTGAGTGTGCTGAGTGTGCTGAGCAGCAGTGTG 3136  
QY 2649 GTGTGTGGACTCCATGCTGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 2708  
Db 3137 GTGTGTGGACTCCATGCTGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 3196  
QY 2709 TAGCATGAGCAGCTGCGCCCTGAAAATTTGTTGAGGCTACTGTACTGTACTGTACTGTACT 2768  
Db 3197 TAGCATGAGCAGCTGCGCCCTGAAAATTTGTTGAGGCTACTGTACTGTACTGTACTGTACT 3256  
QY 2769 GAGCAACAGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2828  
Db 3257 GAGCAACAGGCTGTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3316  
QY 2829 AGAGGGTTCCTTAAAGGACAGTGTGAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2888  
Db 3317 TGAGGAGCAGCTATTAAGGACCTGTGAGTGTGCTTCTTCTTCTTCTTCTTCTTCTTCT 3376  
QY 2889 TCACAGCCTGCTCAATTCAGCATGTTGTGTAGAGGACAGCAGATACACATGTTCTTT 2948

Db 3377 TCACAGCCCCCTTGTAACTCAGCATGTCTAGAGGACAGCATACAACCTGGTCTTT 3436  
Qy 2949 CATTCACTGTCCAGCTTGCATGCAACGGCCACAGTAATGCAATCAATCAGACATCTG 3008  
Db 3437 CATTCACTGTCCAGCTTGCATGCAACGGCCACAGTAATGCAATCAATCAGACATCTG 3496  
Qy 3009 TGAGAAAGTGTAGAACTGTGACCAAGGCAAGCACTGGAGACCTGCATATCTGGCTTCTA 3068  
Db 3497 TGAGAAAGTGTAGAACTGTGACCAAGGCAAGCACTGGAGACCTGCATATCTGGCTTCTA 3556  
Qy 3069 CGGTGATCCCACTGAGGGAATGTCAGCATGCAAGTGAATGCAATGGCGACGGTCTCT 3128  
Db 3557 TGGTGAACCGACTAATGAGGCAATGTGACCATGCAATGCAATGGCGACGATCACT 3616  
Qy 3129 GTGCAACACCAACACGGGCAAGTGTCTGACCAACCAAGGCGTCAAGGGGACGAGTG 3188  
Db 3617 GTGCAACACCAACACGGGCAAGTGTCTGACCAACCAAGGCGTCAAGGGGACGAGTG 3676  
Qy 3189 CCAGCTATGTAGGTAGAAATCGATACCAAGGAACCTCTCAGAGCAACATGTTATTA 3248  
Db 3677 CCAGCTATGTAGGTAGAAATCGATACCAAGGAACCTCTCAGAGCAACATGTTATTA 3736  
Qy 3249 TACTCTTCTATGTACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTA 3308  
Db 3737 TACCCTTCTATGTACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGCTATTA 3796  
Qy 3309 CACAGCTATCAATTTTGGCTACTCTTACGCAACCAACAGGATTTGGACATGTTTCA 3368  
Db 3797 CACAGCCATCAATTTTGGCTACTCTTACGCAACCAACAGGATTTGGACATGTTTCA 3856  
Qy 3369 CAATGCTTCAAGAAATTTCAACCTCAACATCACTTGGGCTGCGATTTCTCAGTGTGAAC 3428  
Db 3857 CAATGCTTCAAGAAATTTCAACCTCAACATCACTTGGGCTGCGATTTCTCAGTGTGAAC 3916  
Qy 3429 CCAGCTGTGAGAAAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTCAAGATAG 3488  
Db 3917 CCAGACTGTGAGAAAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTCAAGATAG 3976  
Qy 3489 TTTCTCTAATGAGAAATTTGATTTTCGCAACCAACCAATATCACTTTCTTTGTTATGT 3548  
Db 3977 CTCTCTTAAGAAATTTGATTTTCGCAACCAACCAATATCACTTTCTTTGTTATGT 4036  
Qy 3549 CAGTAATTTCACTGGCCCATCAAAATTCAGATTCCTTCTCTCAGCAAGCAATTTTAT 3608  
Db 4037 CAGTAATTTCACTGGCCCATCAAAATTCAGATTCCTTCTCTCAGCAAGCAATTTTAT 4096  
Qy 3609 GGACCTGGTACAGTTCCTGCTACTTCTTCAAGTGTTCCTCTCTTCTCTCTGCTGCTG 3668  
Db 4097 GGACCTGGTACAGTTCCTGCTACTTCTTCAAGTGTTCCTCTCTCTGCTGCTGCTGCT 4156  
Qy 3669 TGTCTGTGTTTGAAGATCAAAACAAAGTTGTGGGCTTCCAGACGTAGAGCAACTTCT 3728  
Db 4157 TGCAGTGGTCTGAGATCAAGCAGAGCTGTGGGATCCAGGCGGAGAGCACTTCT 4216  
Qy 3729 TCGAGAGATGCAACAGATGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3788  
Db 4217 TCGGAGATGCAACAGATGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4276  
Qy 3789 AGATGAGGAGCTTCTGATCTTTATTTGGGGAGATTAAGACTGTTCCAAACCCATGTC 3848  
Db 4277 AGATGAGGAGCTTCTGATCTTTATTTGGGGAGATTAAGACTGTTCCAAACCCATGTC 4336  
Qy 3849 ACTGGAGCGGTGTTTGGCAACAAAGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3908  
Db 4337 ACTGGAGCGGTGTTTGGCAACAAAGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4396  
Qy 3909 AGGCTTGGGTGGCATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3968  
Db 4397 AGGCTTGGGTGGCATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4456  
Qy 3969 GGACATTTCTCAGCAGATCCCATAGTGTACAGGAGAGTCAAGAGCGCTGAGAAACCG 4028  
Db 4457 GGACATTTCTCAGCAGATCCCATAGTGTACAGGAGAGTCAAGAGCGCTGAGAAACCG 4516

Qy 4029 GAAGCAGACGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4068  
Db 4517 GAAGCAGACGCCCCCTGCACAGCCTGGGACCTGCATCTGA 4556

## RESULT 4

US-09-245-041-18  
; Sequence 18, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-18

Query Match 50.4%; Score 2051.8; DB 3; Length 2625;  
Best Local Similarity 97.1%; Pred. No. 0;  
Matches 2089; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

Qy 15 GGCGGCAACTAGGCAAGGCTGAGGAGGAGGACGGCGGCGACGGCGGCGCTCGGGGCGAG 74  
Db 234 GGAGCGCGGCGGCTGCGGCTGCGGCGGCGGCTGCGGCTGCGGCGGCGGCGGCGGCGAA 293  
Qy 75 GAGCGGCGGCGGCTGCGGCTGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCT 134  
Db 294 GGAATGTGACGGCGGCTGCGGCTGCGGCGGCGGCTGCGGCGGCGGCGGCGGCGGCGGCT 353  
Qy 135 CTGCGCGCGGCTGCGGCTGCGGCGGCGGCTGCGGCGGCGGCTGCGGCGGCGGCTTCCAGCTAAC 194  
Db 354 CTGCGCGCGGCTGCGGCTGCGGCGGCGGCTGCGGCGGCGGCTTCCAGCTAAC 413  
Qy 195 TGGATCTTCTGGTGTGACAGATGCACTGGAATTAATAAATAAATAAATAAATAAATAAATAAATA 254  
Db 414 TGGATCTTCTGGTGTGACAGATGCACTGGAATTAATAAATAAATAAATAAATAAATAAATA 473  
Qy 255 GTGGCTCATTTGAAGGACAGCAAAATAGAATAATAGACTTCGTTTCAATCATTTTGGCTAC 314  
Db 474 GTGGCTCATTTGAAGGACAGCAAAATAGAATAATAGACTTCGTTTCAATCATTTTGGCTAC 533  
Qy 315 AGAGTGTAGTTGGACCAATTTATGTTTATGATGGGACTCAATTTATGACCGCTAGT 374  
Db 534 AGAGTGTAGTTGGACCAATTTATGTTTATGATGGGACTCAATTTATGACCGCTAGT 593  
Qy 375 TGCTGCATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTTCCTCTGAGGT 434  
Db 594 TGCTGCATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTTCCTCTGAGGT 653  
Qy 435 TGTTGCCACATCAGGTTATGCTTGTGCAATTTTTTTTGTAGTGTGCTTATTAATTTGAC 494  
Db 654 TGTTGCCACATCAGGTTATGCTTGTGCAATTTTTTTTGTAGTGTGCTTATTAATTTGAC 713  
Qy 495 TGGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 554  
Db 714 TGGATTTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 773  
Qy 555 GTGTGAAGATCAGTAATAGCAGCGATCTGTGTAATGTAATGTTCTGAAATCTGGAAGG 614  
Db 774 GTGTGAAGATCAGTAATAGCAGCGATCTGTGTAATGTAATGTTCTGAAATCTGGAAGG 833



QY 615 TGAAGCATGTGACATCTCTCACTGATACAGCAACTGTGGTCTTCTCATCGAGGCATCTG 674  
DB 834 TGAAGCATGTGACATCTCTCACTGATACAGCAACTGTGGTCTTCTCATCGAGGCATCTG 893  
QY 675 CAATTCAAGTGATGTGAGAGGATGCTCTGCTTCTCAGACTGCGAGGGTCTCTGAGTGTTC 734  
DB 894 CAATTCAAGTGATGTGAGAGGATGCTCTGCTTCTCAGACTGCGAGGGTCTCTGAGTGTTC 953  
QY 735 AGTTCCTGACAGCTAAACAGCTATTTGGACTCGAGAGGAATATTTCAACTTAAAGCT 794  
DB 954 AGTTCCTGACAGCTAAACAGCTATTTGGACTCGAGAGGAATATTTCAACTTAAAGCT 1013  
QY 795 CCCCAGAGCATCTCAATAAGCTGTGTCAATGGAACATTTATGTGGTGTGTTGGAGGATA 854  
DB 1014 CCCCAGAGCATCTCAATAAGCTGTGTCAATGGAACATTTATGTGGTGTGTTGGAGGATA 1073  
QY 855 TATGTTCAACCACTCAGATTATAAATGTTCTAGCGTATGACCTTGTCTTCTAGGAGTG 914  
DB 1074 TATGTTCAACCACTCAGATTATAAATGTTCTAGCGTATGACCTTGTCTTCTAGGAGTG 1133  
QY 915 GTTCTCACTAAACCGTCTGTGACAACTGTGTGTTAGATATGTTCTTCTTCTGCAAT 974  
DB 1134 GTTCTCACTAAACCGTCTGTGACAACTGTGTGTTAGATATGTTCTTCTTCTGCAAT 1193  
QY 975 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTTCAACTGGGAATGTGACCAA 1034  
DB 1194 ATACAAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTTCAACTGGGAATGTGACCAA 1253  
QY 1035 TGAGTTGAGAGTTTTTCAATTCATATGAGTATGAGTGTGTTGACCCCTTAAGGCAA 1094  
DB 1254 TGAGTTGAGAGTTTTTCAATTCATATGAGTATGAGTGTGTTGACCCCTTAAGGCAA 1313  
QY 1095 GGAGCAGTATGAGTGTGGGCACTCTGCAACACATTTGTACACTGAAGATGGCCGAGT 1154  
DB 1314 GGAGCAGTATGAGTGTGGGCACTCTGCAACACATTTGTACACTGAAGATGGCCGAGT 1373  
QY 1155 GGTCTGCTGGTCACTTTTGGTCACTGCGCTCTCTATGAGATATATAAGCAATGTGACAGA 1214  
DB 1374 GGTCTGCTGGTCACTTTTGGTCACTGCGCTCTCTATGAGATATATAAGCAATGTGACAGA 1433  
QY 1215 ATATGATTTGGATAAGAACACATGAGTATATTAACACCCAGGCTGCTTGTGCAAGG 1274  
DB 1434 ATATGATTTGGATAAGAACACATGAGTATATTAACACCCAGGCTGCTTGTGCAAGG 1493  
QY 1275 GGGTTACGGCCATPAGCAGTGTTTACGACCATAGGACCGGCTTATAGTTTCAATGGTGG 1334  
DB 1494 GGGTTACGGCCATPAGCAGTGTTTACGACCATAGGACCGGCTTATAGTTTCAATGGTGG 1553  
QY 1335 CTACAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTTACCGATATGATGT 1394  
DB 1554 CTACAGGCTTTTCAAGTCCCAATAGTACCGGCTTGCAGATGATCTTACCGATATGATGT 1613  
QY 1395 GGATACCCAGATGTGGAACCATTTTAAGGACAGCCGATTTTTCCGTTACTTGTGACACAGC 1454  
DB 1614 GGATACCCAGATGTGGAACCATTTTAAGGACAGCCGATTTTTCCGTTACTTGTGACACAGC 1673  
QY 1455 TGTGATAGTGTGGAACCATGCTGTGTGTTTGGGGAAACACACACATATGACATCTAT 1514  
DB 1674 TGTGATAGTGTGGAACCATGCTGTGTGTTTGGGGAAACACACACATATGACATCTAT 1733  
QY 1515 GAGCCATGGCGCAATATGCTTCTTCAATTTTCAATTTGCTTATGCTTATGCTGTGACCG 1574  
DB 1734 GAGCCATGGCGCAATATGCTTCTTCAATTTTCAATTTGCTTATGCTTATGCTGTGACCG 1793  
QY 1575 CTGGTCAGTGTCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGC 1634  
DB 1794 CTGGTCAGTGTCTTCCAGACCTGATCTCCACATGATGTCAACAGATTTGGCCATTCAGC 1853  
QY 1635 AGTCTTACACACAGCAGCAGTATGTTGGTGGTTCATATGCTCTCTCTCAGGGA 1694  
DB 1854 AGTCTTACACACAGCAGCAGTATGTTGGTGGTTCATATGCTCTCTCTCAGGGA 1913

QY 1695 CATCTCTGTATTACCTCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTATTAGC 1754  
DB 1914 CATCTCTGTATTACCTCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCTTGTATTAGC 1973  
QY 1755 AGCAGGACCTGGTATTCGGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGGC 1814  
DB 1974 AGCAGGACCTGGTATTCGGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGGC 2033  
QY 1815 GCTGGCACTGATGAACAAGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCT 1874  
DB 2034 GCTGGCACTGATGAACAAGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCT 2093  
QY 1875 TGACATGACAGATGTACCCAGCACACAGATTGTTTACAGTGCACAGCCAAACCAATGA 1934  
DB 2094 TGACATGACAGATGTACCCAGCACACAGATTGTTTACAGTGCACAGCCAAACCAATGA 2153  
QY 1935 CTGCCACTGTGCAATGACCAATTTGTCTCCAGCAACACAGCTGCTCAGAGGCCAGAT 1994  
DB 2154 CTGCCACTGTGCAATGACCAATTTGTCTCCAGCAACACAGCTGCTCAGAGGCCAGAT 2213  
QY 1995 CTCCATTTTGTAGTATGAGAAATTCGCCCAAGGATAACCTTATCTACTGTAAACAAGAA 2054  
DB 2214 CTCCATTTTGTAGTATGAGAAATTCGCCCAAGGATAACCTTATCTACTGTAAACAAGAA 2273  
QY 2055 GACAGCTGACAGAGCTGTCCCTCGACACAGAACTGCCAGTGGAGGCCCGGAATCAGGA 2114  
DB 2274 GACAGCTGACAGAGCTGTCCCTCGACACAGAACTGCCAGTGGAGGCCCGGAATCAGGA 2333  
QY 2115 GTGCATTTGCCCTGCCGAAATATCTGTGCAATTTGCTGCAATTTGGTTGG 2165  
DB 2334 GTGCATTTGCCCTGCCGAGTAGGCTTGCAGGTCATCTTGGTGTGTGGG 2384

RESULT 5  
US-09-245-041-12  
; Sequence 12: Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 12  
; LENGTH: 6370  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-12

Query Match 45.6%; Score 1853; DB 3; Length 6370;  
Best Local Similarity 96.3%; Pred. No. 0;  
Matches 1980; Conservative 7; Mismatches 47; Indels 23; Gaps 8;  
QY 2035 ATGTACTACTGTAAAGAGACACAGCTGCAGAGCTGTGCCCTGCAAGCAATCTGTCGCTG 2094  
DB 1 ATGTACTACTGTAAAGAGACACAGCTGCAGAGCTGTGCCCTGCAAGCAATCTGTCGCTG 60  
QY 2095 TGGAGCCCGGAATCAGAGTGCATTTGCCCTGCCGAAATATCTGTGCAATGGCTG 2154  
DB 61 TGGAGCCCGGAATCAGAGTGCATTTGCCCTGCCGAAATATCTGTGCAATGGCTG 120  
QY 2155 CATTTGTTGGAACTCATGTTTGAATAATTTACTTCTGCAAGGAGAAATATGACAAATGCT 2214  
DB 121 CATTTGTTGGAACTCATGTTTGAATAATTTACTTCTGCAAGGAGAAATATGACAAATGCT 180





```
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 2419
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-8

Query Match      35.7%; Score 1451.8; DB 3; Length 2419;
Best Local Similarity 89.6%; Pred. No. 0;
Matches 1561; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 189 ACTAACTGGATCTCTCGGTTTGTGACAGATGGACCTGGAAATATAAATAAAGAA 248
DB 677 ACTAACTGGCTCTCTCGGTTTGTGACAGATGGACCTGGAAATATAAATAAAGAA 736
QY 249 GTGACAGTGGCTCAATTGAAGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTT 308
DB 737 GTGACAGTGGCTCAATTGAAGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTT 796
QY 309 TGCTACAGATGTAGTTGGACCAATTTATATGTTTATGATGGGAACTCAATTTATGCACC 368
DB 797 TGCTACAGATGTAGTTGGACCAATTTATATGTTTATGATGGGAACTCAATTTATGCACC 856
QY 369 GCTAGTTGCTGCAATTTAGTGGCTCAATTTCTGAGAGAGATGGCAATGAGACTGTCCC 428
DB 857 TCTGATTTGCTGCTTTAGTGGCTCAATTTCTGAGAGAGATGGCAATGAGACTGTCCC 916
QY 429 TGAGTTTGTGCCACATCAGTTTATGCTTGTGCAATTTTATGATGCTGCTTTATAA 488
DB 917 TGAGTTTGTGCCACATCAGTTTATGCTTGTGCAATTTTATGATGCTGCTTTATAA 976
QY 489 TTTGACTGGATTTAAATATTACTTACAGTTTGTATGTTGTCNAATTAACCTCAGGCCG 548
DB 977 TCTGACTGGATTTAAATATTACTTACAGTTTGTATGTTGTCNAATTAACCTCAGGCCG 1036
QY 549 AGGAGAGTGTAAAGATCAGTAATAGCAGCGATCTGTTGAATGTGAATGTTCTGAAAACTG 608
DB 1037 AGGAGAGTGTAAAGATCAGTAATAGCAGCGATCTGTTGAATGTGAATGTTCTGAAAACTG 1096
QY 609 GAAAGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 668
DB 1097 GAAAGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCCTCATCGAGG 1156
QY 669 CATCTGCAATTTCAAGTGTGTACAGAGATGCTCCTCTCTCAGACTGCGAGGTCCTGG 728
DB 1157 CATCTGTAATGAAGCATACAGAGGTGCTCCTCTCTCAGACTGCGAGGTCCTGG 1216
QY 729 ATGTTCAAGTTCCTGTACCACTAACCACTCAATTTTGGACTCGAGAGGAATATTCATCTT 788
DB 1217 ATGTTCAAGTTCCTGTGCGCACTAACCACTCAATTTTGGACTCGAGAGGAATATTCATCTT 1276
QY 789 AAGCTCTCCAGAGCATCTCAATAAGCTGTGCTCAATGGAACANTATGTGGTTGTTGG 848
DB 1277 AAGCTCTCCAGAGCATCTCAATAAGCTGTGCTCAATGGAACANTATGTGGTTGTTGG 1336
QY 849 AGGATATATGTTCAACCACTCAGATTTAATACATGTTCTAGCGTATGACCTTGTCTTAG 908
DB 1337 CGGATATATGTTCAACCACTCAGATTTAATACATGTTCTAGCGTATGACCTTGTCTTAG 1396
QY 909 GGAGTGGCTCCACTAAACGGTTCTGTGAACATGTGGTTGTAGATGTTGTCATCTTT 968
DB 1397 GGAGTGGCTCCACTAAACGGTTCTGTGAACATGTGGTTGTAGATGTTGTCATCTTT 1456
QY 969 GGCATTATACAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTTCACTGGGAATGT 1028
DB 1457 GGCATTATACAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTTCAACAGGGAAGT 1516
QY 1029 GACCAATGAGTTGAGAGTTTTCATCATTAATGAGTCACTGGGTGTTGTGACCCCTAA 1089
DB 1517 GACCAATGAGTTGAGAGTTTTCATCATTAATGAGTCACTGGGTGTTGTGACCCCTAA 1576
QY 1089 GGCAGAGGAGCAGTATGCAAGTGGTTGGGCACTCTGCACACATTTGTACACTGAAGAATGG 1148
```

```
DB 1577 AGCTAAGGATCAGTATGAGTGGTTGGACACTCAGCACACATTTGTACACTGGCATCGG 1636
QY 1149 CCGAGTGTGTCATGCTGGTTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCAATGT 1208
DB 1637 CCGTGTGTCATGTTGGTTCATCTTTGGTCACTGCCCTCTCTATGGATATATAAGCGTTGT 1696
QY 1209 GCAGGAATATGATTTGGATAAGAAACACATGAGATATATTACACACCCAGGGTGCCTTGT 1268
DB 1697 GCAGGAATATGATTTGGATAAGAAACACATGAGATATATTACATACTCAGGGTGCCTTGT 1756
QY 1269 GCAGGGGGTTACGGCCATAGCAGTGTTCAGGACCATAGGACCGAGCCCTATACGTTCA 1328
DB 1757 GCAGGGGGTTATGGCCACAGTAGTGTTCATGATGACAGGACCAAGGCTCTGACGTTCA 1816
QY 1329 TGGTGGCTTACAGGCTTTTCAGTGCCTAATAGTACGGCTTCGAGATGATCTCTACCGATA 1388
DB 1817 TGGTGGCTTACAGGCTTTTCAGGCTTCGAGGCTTCGAGATGATCTCTACGATA 1876
QY 1389 TGATGTGATATCCAGATGTGGACCAATCTTAAGGACAGCGGATTTTCGTTACTTGCA 1448
DB 1877 CGATGTGATATCTCAGATGTGGACCAATCTTAAGGACAGCGGATTTTCGTTACTTGCA 1936
QY 1449 CACAGCTGTGATAGTGTGGAACCAATGCTGTGTTTGGGGGAAACACACACATGACAC 1508
DB 1937 TACAGCTGTGATAGTGTGGAACCAATGCTGTGTTTGGAGGNAACACACACATGACAC 1996
QY 1509 ATCTATGAGGCAATGGCCAAATGCTTCTCTCAGATTTTATGCGCTTATGACATTTGCTG 1568
DB 1997 TCCATGAGCCACCGTGCCTTCTCTCTCGGACTTTCATGGCTTATGACATTTGCTG 2056
QY 1569 TGACCGCTGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGSCCA 1628
DB 2057 TGACCGATGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGSCCA 2116
QY 1629 TTAGCAGTCTTTACACAAACGACCATGATGTGTTTGGTGGTTTCAATAGTCTCCTCT 1688
DB 2117 TTAGCAGTCTTTTACAAACGACCATGATGTGTTTGGTGGTTTCAACAGCTCCTCT 2176
QY 1689 CAGGACATCTGTTTACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCGCTTG 1748
DB 2177 CAGTGAAGTCTTGTGTTTACCTCGGAGCAGTGTGATGCAACCGAGTGAAGCTGCTTG 2236
QY 1749 TTTAGCAGCAGGACTGTTTGGTGTGTGGAACACAGGCTGCTTCTCAGTGTATCTC 1808
DB 2237 TGTGCGACAGGACTGTTTGGTGTGTGGAACACAGTGTGATGCTTCTCGATGTACCTC 2296
QY 1809 GTGGGCTCGGCACTGATGAAAGAAAGTTAAATCAGATGTTTTCGAAAAG 1868
DB 2297 CTGGAGTTGGCACTGAAGAACAGCAAAAAAGTTAAAAATCAGAGTGTTCCTAAAAG 2356
QY 1869 AACTCTTGACATCAGATGTGACAGCACAGATTTGTACAGCTGACAGCCAAACAC 1928
DB 2357 AACTCTTGACATCAGATGTGACAGCACAGATTTGTACAGCTGACAGCCAAACAC 2416
QY 1929 CAA 1931
DB 2417 CAA 2419
```

## RESULT 7

```
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245, 041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
```

```

; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
; US-09-245-041-10

Query Match      11.4%; Score 463; DB 3; Length 1051;
Best Local Similarity 85.2%; Pred. No. 4e-127;
Matches 517; Conservative 0; Mismatches 90; Indels 0; Gaps 0;

Qy 15 GCGCGCAACTGAGGCAAGCGCTGAGGAGGAGCAGCGCGCGCAGCGCAGCGCTCCGCGGCAG 74
Db 441 GGAAGCGCGAGCGCGCGCTGCGGTGCGCGCGCGCGGTGTCGGGTGCGCGCGCAGCGCGAGGCCAA 500
Qy 75 GAGCGCGCGCGCGCAGCTGTGTCAACGCGCGGTGCGTGCACACCTGTGGCACCAGCGCGCAGCTGCGT 134
Db 501 GGAATGTGACCGGCGCGTGTGTCAACGCGCGCGCGCTGCAACCCCTGGCACCCGCGCCAGTGGCT 560
Qy 135 CTGCCCCCGCGGTGGGTGGCGGAGCAATGCCAGACA CTGCGGGGGCGCGCTTCAGACTAAC 194
Db 561 CTGCCCCACGCGGTGGGTGGCGGAGCAATGCCAGACA CTGCGGGGGCGCGCTTCAGACTAAC 620
Qy 195 TGGATCTTCTCGGGTTCTGTGACAGATGACCTGGAAATTTATAAATACAAAAACGAAGTGCAC 254
Db 621 TGGCTCTTCTGGATTCTGTAAACAGATGACCTGGGAATTTATAATATAAGACGAAGTGCAC 680
Qy 255 GTGGCTCAITGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAATCATTTTGTCTAC 314
Db 681 ATGCTCATTTGAAGGACAGCCAAATAGAATAATGAGACTTCGTTTCAACCATTTTGTCTAC 740
Qy 315 AGACTGTAGTTGGACCAATTTATATGTTTATGATGGGACTCAATTTTATGCACGGCTAGT 374
Db 741 AGAATGTAGCTGGACCAATTTATATGTTTATGATGGGACTCAATCTACGCACCTCTGAT 800
Qy 375 TGCTGCATTTAGTGGCGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGT 434
Db 801 TGCTGCCTTTAGTGGCGCTCATTTGTTCTGAAAGAGATGGCAATGAGACGGCTCCTGAGGT 860
Qy 435 TGTGGCACATCAGGTTATGCTTGTGCAATTTTATGATGATGCTGCTGTTATATATTTGAC 494
Db 861 CACTGTCACTTCAGGTTATGCACTGCTGCATTTTTTTCAGTGATGCTGCTTATATCTGAC 920
Qy 495 TGGATTTAAATATTACTTACAGTTTGATGTGTCCTGCAATTAATGCTCAGCGCGAGGAGA 554
Db 921 TGGATTTAAATATCACTTACAAATTTTGACATGTGTCGGAATTAATTTGCTCAGCGCGAGGAGA 980
Qy 555 GTGTAAGATCAGTAATAGCAGCGATACGTGTTGAATGTGAATGTGTTCTGAAAACTCGAAAGG 614
Db 981 GTGTAAGAGCAGTAACAGCAGCGCGCTGTTTGTGAGTGTGAATGTGTTCTGAAAACTCGAAAGG 1040
Qy 615 TGAAGCA 621
Db 1041 GCGCGGA 1047

RESULT 8
US-09-702-705-966/c
; Sequence 966, Application US/09702705
; Patent No. 6504010
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedwick, Tom
; APPLICANT: Carter, Darick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun

```

	/	TITLE OF INVENTION:	COMPOSITIONS AND METHODS FOR THE THERAPY AND	
	/	TITLE OF INVENTION:	DIAGNOSIS OF LUNG CANCER	
	/	FILE REFERENCE:	210121.478C14	
	/	CURRENT APPLICATION NUMBER:	US/09/702,705	
	/	CURRENT FILING DATE:	2000-10-30	
	/	NUMBER OF SEQ ID NOS:	1833	
	/	SOFTWARE:	FastSeq for Windows Version 3.0	
	/	SEQ ID NO	966	
	/	LENGTH:	246	
	/	TYPE:	DNA	
	/	ORGANISM:	Homo sapien	
	/	US-09-702-705-966		
		Query Match	6.0%; Score 246; DB 4; Length 246;	
		Best Local Similarity	100.0%; Pred. No. 5.2e-63;	
		Matches 246; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy		2293	CTCGGAATGTCAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGGC	2352
Db		246	CTCGGAATGTCAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGGC	187
Qy		2353	CTTCGGAAGATCAATGTGTCCTACTGCTGCTGGGAAGATAATGCCCCAATTTCACAATAGT	2412
Db		186	CTTCGGAAGATCAATGTGTCCTACTGCTGCTGGGAAGATAATGCCCCAATTTCACAATAGT	127
Qy		2413	TTACTACAGTGAGTGGCGTCTGAGGCCAGTGATGCTGGATTCCTGTGGAATTTTATCAGAA	2472
Db		126	TTACTACAGTGAGTGGCGTCTGAGGCCAGTGATGCTGGATTCCTGTGGAATTTTATCAGAA	67
Qy		2473	CCAGTACTCGGGAGCTGGAAGCTGCAACCCTGCATCAACCCATCAATGTAGTGTCTGT	2532
Db		66	CCAGTACTCGGGAGCTGGAAGCTGCAACCCTGCATCAACCCATCAATGTAGTGTCTGT	7
Qy		2533	GAAAGG 2538	
Db		6	GAAAGG 1	
		RESULT 9		
		US-09-736-457-966/c		
		Sequence 966, Application US/09736457		
		Patent No. 6509448		
		GENERAL INFORMATION:		
		APPLICANT:	Wang, Tongtong	
		APPLICANT:	Bangur, Chaitanya S.	
		APPLICANT:	Lodes, Michael A.	
		APPLICANT:	Fanger, Gary	
		APPLICANT:	Vedvick, Tom	
		APPLICANT:	Carter, Darrick	
		APPLICANT:	Retter, Marc	
		APPLICANT:	Mannion, Jane	
		APPLICANT:	Fan, Liqun	
		APPLICANT:	Wang, AiJun	
		TITLE OF INVENTION:	COMPOSITIONS AND METHODS FOR THE THERAPY AND	
		TITLE OF INVENTION:	DIAGNOSIS OF LUNG CANCER	
		FILE REFERENCE:	210121.478C15	
		CURRENT APPLICATION NUMBER:	US/09/736,457	
		CURRENT FILING DATE:	2000-12-13	
		NUMBER OF SEQ ID NOS:	1864	
		SOFTWARE:	FastSeq for Windows Version 3.0	
		SEQ ID NO	966	
		LENGTH:	246	
		TYPE:	DNA	
		ORGANISM:	Homo sapien	
		US-09-736-457-966		
		Query Match	6.0%; Score 246; DB 4; Length 246;	
		Best Local Similarity	100.0%; Pred. No. 5.2e-63;	
		Matches 246; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
Qy		2293	CTCGGAATGTCAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGGC	2352
Db		246	CTCGGAATGTCAGTCACTCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGGC	187

QY 2353 CTTGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGTCCTCAATTTACAATAGT 2412  
 DB 186 CTTGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGTCCTCAATTTACAATAGT 127  
 QY 2413 TTACTACAGTGGATGCGCTGTGAGCCAGTGCCTGGATTCCTGTGGAAATTTATCAGAA 2472  
 DB 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGCCTGGATTCCTGTGGAAATTTATCAGAA 67  
 QY 2473 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532  
 DB 66 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7  
 QY 2533 GAAAGG 2538  
 DB 6 GAAAGG 1  
 RESULT 10  
 US-09-614-124B-966/c  
 ; Sequence 966, Application US/09614124B  
 ; Patent No. 8630574  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Lodes, Michael A.  
 ; APPLICANT: Fanger, Gary  
 ; APPLICANT: Vedvick, Tom  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Fan, Liqun  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND  
 ; FILE REFERENCE: 210121.478C9  
 ; CURRENT FILING DATE: 2001-07-11  
 ; NUMBER OF SEQ ID NOS: 1668  
 ; SOFTWARE: Fast-SEQ for Windows Version 3.0  
 ; SEQ ID NO 966  
 ; LENGTH: 246  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-09-614-124B-966

Query Match 6.0%; Score 246; DB 4; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-63;  
 Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2293 CTGCGAATATGAGTCATCTCAGACATGTCCAAGCTCACTTAACCCCATGGGTGCGC 2352  
 DB 246 CTGCGAATATGAGTCATCTCAGACATGTCCAAGCTCACTTAACCCCATGGGTGCGC 187  
 QY 2353 CTTGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGTCCTCAATTTACAATAGT 2412  
 DB 186 CTTGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGTCCTCAATTTACAATAGT 127  
 QY 2413 TTACTACAGTGGATGCGCTGTGAGCCAGTGCCTGGATTCCTGTGGAAATTTATCAGAA 2472  
 DB 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGCCTGGATTCCTGTGGAAATTTATCAGAA 67  
 QY 2473 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532  
 DB 66 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7  
 QY 2533 GAAAGG 2538  
 DB 6 GAAAGG 1

RESULT 11  
 US-09-671-325-966/c  
 ; Sequence 966, Application US/09671325  
 ; Patent No. 6667154

; GENERAL INFORMATION:  
 ; APPLICANT: Wang, Tongtong  
 ; APPLICANT: Bangur, Chaitanya S.  
 ; APPLICANT: Lodes, Michael A.  
 ; APPLICANT: Fanger, Gary  
 ; APPLICANT: Vedvick, Tom  
 ; APPLICANT: Carter, Darrick  
 ; APPLICANT: Retter, Marc  
 ; APPLICANT: Mannion, Jane  
 ; APPLICANT: Fan, Liqun  
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
 ; FILE REFERENCE: 210121.478C12  
 ; CURRENT FILING DATE: 2000-09-26  
 ; NUMBER OF SEQ ID NOS: 1825  
 ; SOFTWARE: Fast-SEQ for Windows Version 3.0  
 ; SEQ ID NO 966  
 ; LENGTH: 246  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapien  
 US-09-671-325-966

Query Match 6.0%; Score 246; DB 4; Length 246;  
 Best Local Similarity 100.0%; Pred. No. 5.2e-63;  
 Matches 246; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2293 CTGCGAATATGAGTCATCTCAGACATGTCCAAGCTCACTTAACCCCATGGGTGCGC 2352  
 DB 246 CTGCGAATATGAGTCATCTCAGACATGTCCAAGCTCACTTAACCCCATGGGTGCGC 187  
 QY 2353 CTTGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGTCCTCAATTTACAATAGT 2412  
 DB 186 CTTGGAAGATCAATGTGCTCTACTGTGTGGGAAGATATGTCCTCAATTTACAATAGT 127  
 QY 2413 TTACTACAGTGGATGCGCTGTGAGCCAGTGCCTGGATTCCTGTGGAAATTTATCAGAA 2472  
 DB 126 TTACTACAGTGGATGCGCTGTGAGCCAGTGCCTGGATTCCTGTGGAAATTTATCAGAA 67  
 QY 2473 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 2532  
 DB 66 CCCAGTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7  
 QY 2533 GAAAGG 2538  
 DB 6 GAAAGG 1

RESULT 12  
 US-09-833-381-1916  
 ; Sequence 1916, Application US/09833381  
 ; Patent No. 6672186  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Robison, Keith E.  
 ; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs  
 ; FILE REFERENCE: 5800-119  
 ; CURRENT APPLICATION NUMBER: US/09/833,381  
 ; CURRENT FILING DATE: 2001-04-11  
 ; PRIOR APPLICATION NUMBER: 09/516,448  
 ; PRIOR FILING DATE: 2000-02-29  
 ; NUMBER OF SEQ ID NOS: 2050  
 ; SOFTWARE: Fast-SEQ for Windows Version 3.0  
 ; SEQ ID NO 1916  
 ; LENGTH: 415  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; FEATURE:  
 ; NAME/KEY: misc feature  
 ; LOCATION: (1)-(415)  
 ; OTHER INFORMATION: n = A,T,C or G  
 US-09-833-381-1916

Query Match 5.2%; Score 212.8; DB 4; Length 415;

	Best Local Similarity	73.6%;	Pred.	No. 5.8e-53;	Mismatches	0;	Gaps	0;
	Matches	271;	Conservative	0;	Mismatches	97;	Indels	0;
QY	200	CTTCTGGGTTTTGTACAGATGCACCTGGAAATTATAAATCAAAAAGAAGTCACGTGGC	259					
Dd	5	CGTCGGGATATTTAACAGATGCCCAAATTAACTATAAATAAATAACTAAATGTAAGTGCTGGC	64					
QY	260	TCAITGAAGGACGAGCCAAATAGAAATATAGACACTTGGTTTCANTCATTTTGGTCTACAGAGT	319					
Dd	65	TCAITTCGAAGGCTATCCAATCACGTTGTTAAGATTAAAGATTCATCATTTTGTCTACAGAAT	124					
QY	320	GTAGTTGGGACCAITTATATGCTTTATGATGGGACTCAATTTATGACCCTAGTTGTCTG	379					
Dd	125	GTAGCTGGGATCANATGTATGTTTATGATGAGAAATCAATATATGACCCTTTAATAGCTG	184					
QY	380	CATTTAGTGGCCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGCTCCCTGAGGTTGTG	439					
Dd	185	TACTTAGTGGTTTTGATGTCCTGAAATAGGGGCAATGAACCTGCTGAGTTGTTA	244					
QY	440	CCACATCAGGTTATGCCTTGCTGCATTTTTTTAGTGATGCTGCTTTAAATTTGACTGGAT	499					
Dd	245	CTACATCTGGCTATGCACCTGTACATTTTTTTAGTGATGCTGGGTATATCTAACTGGTT	304					
QY	500	TTAATAATTACTTACAGTTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGTA	559					
Dd	305	TCAACAATTTTCTATTCAATCAATCTTGTCTCTCAAAATTTGCTGTGTCATGGGAAGTGTA	364					
QY	560	AGATCAGT	567					
Dd	365	CAACTAGT	372					

```

RESULT 13
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF OBESITY
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5

Query Match          5.0%; Score 202.8; DB 3; Length 90050;
Best Local Similarity 81.8%; Pred. No. 2.8e-48;
Matches 234; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY      947 TTGTTAGATATGGTCATCTTTGGCAATTATACAAGGATAAAATTTACATGTATGGAGAA 1006
DB      6931 TGGTTTAAAAATCGTTTACATCTTTCTTATAGGATAAAATCTACATGTATGGAGAA 6990

QY      1007 AAATTGANTCAACTGGGAATGTGACCAATGATGTAGAGATTTTTCACATTCATAATCAGT 1066
DB      6991 AAATTGANTCAACAGGGAACGTGACCAATGATGTAGAGATTTTTCATAATTCATAATGAAT 7050

QY      1067 CATGGGTGTGTGTGACCCCTAAGGCAAAAGGAGCAGTATGCAGTGGTTGGGCACCTCTGCAC 1126
DB      7051 CATGGGTATGTGTTAACTCGAAAGCTAAGGATCAGTATGCAGTGGTTGGACACTCAGCAC 7110

QY      1127 ACATTTGTTACCTGAAGAATGCGCGAGTGGTCATGCTGGTCATCTTTGGTCACTGGCCCTC 1186

```

```

Db      7111 ACATTGTTACACTGGGCATCTGGCCGTGGTGTGATGTTGGTCATCTCTGGTCATTGCCAC 7170
Qy      1187 TCTATGGATATATAAGCAATGTCAGAGAAATATGATTGGATAAGAA 1232
Db      7171 TCTATGGATATATAAGCGTGTGTCAGAGAAATGACATTGGTATGTA 7216

RESULT 14
US-09-245-041-3/c
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 17056
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-3

Query Match      4.2%; Score 169.2; DB 3; Length 17056;
Best Local Similarity 86.9%; Pred. No. 8.3e-39;
Matches 186; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

Qy      510 TTACAGTTTTCATATGTCGAAATAACTGCTCAGCGCAGGAGAGTGTAAAGATCAGTAA 569
Db      4090 TTGTTAGTTTTCACATGTGTCGAAATAATTTGCTCAGCGCAGGAGAGTGTAAAGCAGTAA 4031
Qy      570 TAGCAGCGATACTGTTGAATGTGAATGTTCTTGAAACTGGAAGCTGAAGCATGTGACAT 629
Db      4030 CAGCAGCAGCGCTGTTGAGTGTGAATGTTCTTGAAACTTGAAGAGGGAGTCTGTGTGACAT 3971
Qy      630 TCCTCACTGTACAGACAACTGTGGTGTTCCTTCATCGAGGCATCTGCAATTCGAAGTGATGT 689
Db      3970 TCCTCACTGTACAGACAACTGTGGCTTTCCTCACCAGGCATCTGTAATGCAAGCGATAC 3911
Qy      690 CAGAGGATGTCCTCGTTCTCTCAGACTGGCAGGGT 723
Db      3910 CAGAGGGTGCTCCTGCTTTTCCTCTCACTGGCAGGGT 3877

```

```

RESULT 15
US-09-041-6/c
; Sequence 6, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 226

```

```
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-6

Query Match      3.2%; Score 132.2; DB 3; Length 226;
Best Local Similarity 78.6%; Pred. No. 3.6e-29;
Matches 158; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 2389 GATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATCCGTCGTAGCCGCCAGTGATGCT 2448
Db      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 225  GATATGTCCTCCATTCACAAATAGTTTCTGTCAGTGGATCCATCTGAGCCCGAGTGATGCT 166
Db      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2449 GGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATC 2508
Db      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 165  GGCTTCTGTGGGATCTTGTTCAGAGCCTAGTACTCGGGGATTAAGGCTGCAACCTGCATC 106
Db      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2509 AACCCACTCAATGGTAGTGTCTGTGAAGCCCTGCAACCAAGTGTCTAAGCAGTCCCGG 2568
Db      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 105  AACCTCTCAATGGCAGCGTCTGTGAAGGCTGTGAAGGACATGGGTGCATATAGTGCT 46
Db      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 2569 ACACCATGTCCCTTGAGGACA 2589
Db      ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
QY 45  CCAGGAGGAGCCAGACAGCA 25
```

Search completed: March 5, 2004, 12:20:15  
Job time : 192.783 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 29, 2004, 14:49:01 ; Search time 11557 Seconds  
(without alignments)  
5359.280 Million cell updates/sec

Title: US-09-787-097-12  
Perfect score: 7915  
Sequence: 1 MVAAATATEALRRRTAATA.....KSGAVRNEKQPPAQPGTCI 1429

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q/cgn2\_1/USPTO\_spool\_p/US09787097/runat\_26022004\_062934\_23420/app\_query.fasta\_1.1607  
-DB=GenEmbl -QFWT=fastap -SUFFIX=rgc -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi -LIST=45  
-DOALIGN=200 -THR SCORE=500 -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09787097@cgn 1 1 7794 -runat 26022004 062934 23420 -NCPU=6 -ICPU=3  
-NO MWAP -LARGQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG  
-DEV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

1: gb.ba:\*  
2: gb.htg:\*  
3: gb.in:\*  
4: gb.om:\*  
5: gb.ov:\*  
6: gb.pat:\*  
7: gb.ph:\*  
8: gb.pl:\*  
9: gb.pr:\*  
10: gb.ro:\*  
11: gb.sts:\*  
12: gb.sy:\*  
13: gb.un:\*  
14: gb.vi:\*  
15: em.ba:\*  
16: em.fun:\*  
17: em.hum:\*  
18: em.in:\*  
19: em.mu:\*  
20: em.mu:\*  
21: em.or:\*  
22: em.ov:\*  
23: em.pat:\*  
24: em.ph:\*  
25: em.pl:\*  
26: em.ro:\*  
27: em.sts:\*  
28: em.un:\*

29: em.vi:\*  
30: em.htg\_hum:\*  
31: em.htg\_inv:\*  
32: em.htg\_other:\*  
33: em.htg\_mus:\*  
34: em.htg\_pln:\*  
35: em.htg\_rdt:\*  
36: em.htg\_mam:\*  
37: em.htg\_vit:\*  
38: em.sy:\*  
39: em.htgo\_hum:\*  
40: em.htgo\_mus:\*  
41: em.htgo\_other:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	7696.5	97.2	8599	6	AR164816	AR164816 Sequence
2	7405	93.6	4493	4	AF531101	AF531101 Bos tauri
3	7386	93.3	4284	10	AB062913	AB062913 Mesocric
4	7346.5	92.8	6083	10	AF116897	AF116897 Mus muscu
5	7342.5	92.8	8739	10	AB038387	AB038387 Rattus no
6	7303	92.3	8827	6	AR164807	AR164807 Sequence
7	7302.5	92.3	4313	10	AF119821	AF119821 Mus muscu
8	7089	89.6	3819	9	AF106861	AF106861 Homo sapi
9	6870.5	86.8	4072	6	AR164817	AR164817 Sequence
10	6612	83.5	3597	9	AF034957	AF034957 Homo sapi
11	6509.5	82.2	4500	10	AB038388	AB038388 Rattus no
12	5624	71.1	5683	9	AB093245	AB093245 Mus muscu
13	4638.5	58.6	4140	9	AY442317	AY442317 Homo sapi
14	4622	58.4	4463	9	AK127277	AK127277 Homo sapi
15	4136.5	52.3	2625	6	AR164818	AR164818 Sequence
16	3880	49.0	3609	6	AX503776	AX503776 Sequence
17	3632.5	45.9	7302	9	AB011106	AB011106 Homo sapi
18	3491.5	44.1	6370	6	AR164815	AR164815 Sequence
19	3435	43.4	2419	6	AR164813	AR164813 Sequence
20	3377.5	42.7	5002	10	BC027764	BC027764 Mus muscu
21	3145	39.7	2578	9	AK000356	AK000356 Homo sapi
22	3102.5	39.2	4997	10	BC050020	BC050020 Mus muscu
23	2495	31.5	5632	9	AB011120	AB011120 Homo sapi
24	2328	29.4	5271	3	AY051898	AY051898 Drosophil
25	2032.5	25.7	180499	3	AC007817	AC007817 Drosophil
26	2032.5	25.7	224910	3	AE003761	AE003761 Drosophil
27	1991	25.2	14345	2	AC013047	AC013047 Drosophil
28	1939	24.5	3990	3	AF339882	AF339882 Caenorhab
29	1825	23.1	1012	6	E26747	E26747 Structure a
30	1797.5	22.7	3667	3	BT011125	BT011125 Drosophil
31	1529.5	19.3	3297	10	BC030872	BC030872 Mus muscu
32	1302.5	16.5	25305	3	CEP33C8	CEP33C8
33	1290	16.3	1818	9	BC047716	BC047716 Homo sapi
34	1226.5	15.5	1051	6	AR164814	AR164814 Sequence
35	1014	12.8	543	6	AX778545	AX778545 Sequence
36	977.5	12.3	8875	6	BD185470	BD185470 Membrane
37	977.5	12.3	9921	6	BD185469	BD185469 Membrane
38	960.5	12.1	8589	6	AX921022	AX921022 Sequence
39	950	12.0	6056	6	AX592784	AX592784 Sequence
40	923	11.7	8307	9	AY280362	AY280362 Homo sapi
41	922	11.6	6228	9	AB011541	AB011541 Homo sapi
42	895	11.3	169427	3	AC007416	AC007416 Drosophil
43	895	11.3	289516	3	AE003619	AE003619 Drosophil
44	869.5	11.0	36993	2	AC018007	AC018007 Drosophil
45	841.5	10.6	5693	10	BC036727	BC036727 Mus muscu

# ALIGNMENTS

RESULT 1





577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596  
1726 ACACTATATGACCATGGCCAAATGCTTCTCTTTCAGATTTTCATGGCTATGACATGGC 1785  
597 CysAspArgTTPSerValLeuProArgProAspLeuHisAspValAsnArgPheGly 616  
1786 TGTGACCGCTGGTCAGTGCCTTCCAGACCTGATCTCCACCATGATGTCACAGATTTGGC 1845  
617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
1846 CATTACAGAGCTTACACAACAGCACCATGATATGTTGGTGGTTCCTCAATAGTCTCCTC 1905  
637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
1906 CTCAGGCATCTCTGGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCT 1965  
657 CysLeuAlaAlaGlyProGlyIleArgCysValTTPAsnThrGlySerSerGlnCysIle 676  
1966 TGTTTAGCAGGAGGACCTGGTATTCGGTGTGTGGGAACACAGGGTGTCTCAGTGTATC 2025  
677 SerTTPAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696  
2026 TCGTGGCGCTGGCAACTGATGAACAGAGAAAGATTAAATCAGAAATGTTTCCAAA 2085  
697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
2086 AGAACTCTTGACCATGACAGATGTGACAGACACAGATTTGTACAGCTGCACAGCCAAC 2145  
717 ThrAsnAspCysHisTTPCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
2146 ACCAATGATGCCACTGGTGTGAATGACCATTTGTGCCAGGACCAACAGCTGCTCAGAA 2205  
737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
2206 GGCCAGATCTCCATTTTAGGTATGAGAAATGGCCCAAGGATAACCCCTATGTAAGTACTGT 2265  
757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTTPGluProArg 776  
2266 AACAGAAAGACCATGTCAGGAGCTGTGCCCTGGACAGAACTGCCAGTGGAGGCCCGG 2325  
777 AsnGlnGlyCysIleAlaLeuProGluAsnIleCysGlyIleGlyTTPHisLeuValGly 796  
2326 AATCAGGAGTGCATTTGCCCTGCCGAAATATCTGTGCAATGGCTGGCAATTTGGTGA 2385  
797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
2386 AACTCATGTTTGAATTAATTAATCTGCAAGGAAATATGCAATGCTTAATTTGTTCTGT 2445  
817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836  
2446 AGGAACCAATGGCCCTTTTGGCTTCTTACAAACCAAGAGAGTATGATTTGTCCTT 2505  
837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTTP 856  
2506 AAGCAGCTCGAATAATGAGTCACTCTCAGACATGTCCAAGCTCACCTTAACCCCATGG 2565  
857 ValGlyLeuArgLysIleAsnValSerTyrTTPCysTTPGluAspMetSerProPheThr 876  
2566 GTCCGCTTTCGAAGATCAATGTGTCCTACTGGTGTGGGAAGATATGTCCTCCATTTACA 2625  
877 AsnSerLeuLeuGlnTTPMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
2626 AATAGTTTACTACAGTGGATGGCTGTGAGCCAGTGTGCTGGATTTCTGTGGAAATTTTA 2685  
897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
2686 TCAGAACCCAGTACTCGGGACTGAAGGCTGCAACCTGATCAACCCACTCAATGGTAGT 2745  
917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
2746 GTCTGTGAAGGCTGCARACACACAGTGTAGCAGTGCAGACACCATGTCCTTGGG 2805  
937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTTPCysSerAsnMetLys 956

2806 ACAGCATGTGGAGATTGCCACGCGCACCTCTGAGTGCATGTGGTGCAGCAACATGAAG 2865  
957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTTP 976  
2866 CAGTGTGTGAGCTCCAAATGCCCTATGTGGCTCTCTCTCTTTGGCCAGTGTATGAATGG 2925  
977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
2926 TATACGATGAGCACCTGCCCCCTGAAATTTGTTCAGGTACTGTACCTGTATGATCAATGC 9985  
997 LeuGluGlnProGlyCysGlyTTPCysThrAspProSerAsnThrGlyLysGlyLysCys 1016  
2986 TTCGAGCAACAGGCTGTGGCTGTACTGATCCCGCAATACTGGCAAGGGAAATGC 3045  
1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036  
3046 ATAGAGGGTTCTTATAAGACCCAGTGAAGATGCTTCGCAAGCCCTACAGGAATTTTC 3105  
1037 TyrProGlnProLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTTPSer 1056  
3106 TATCCACAGCCCTGCTCAATTCACGATGTGTCTAGAGACAGCAGATACAACTGCTCT 3165  
1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
3166 TTCATTCATGTCAGCTTGCCAAATGCAAGCGCCACAGTAAATGCATCAATCAGAGCATC 3225  
1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096  
3226 TGTGAGAAGTGTGAGAACCTTGACACAGGCAAGCACTGGCAGACCTGCAATCTGGCTTC 3285  
1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116  
3286 TAGGTGTATCCCAACCAATGGAGGAAATGTCCAGCATGCAATGGCAGCGGTCT 3345  
1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136  
3346 CTGTGCAACACCAACACGGGCAAGTGTCTTGCAACCAAGGCGCTCAAGGGGAGCAG 3405  
1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156  
3406 TGGCAGCTATGTGAGGTAGAAATCGAATACCAAGGAAACCCCTCTCAGAGGAACATGTTAT 3465  
1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176  
3466 TATACTCTTCTTATGACTATCAGTTTACCTTTAGTCTATCCCGAAGATGATCGCTAT 3525  
1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196  
3526 TACACAGCTATCAATTTTGTGGCTACTCTCTGACGAAACAAACAGGATTTGGACATGTTTC 3585  
1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTTPAlaAlaSerPheSerAlaGly 1216  
3586 ATCAATGCCCTCCAGAAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGA 3645  
1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236  
3646 ACCCAGGCTGCAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTCAAAAGAT 3705  
1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256  
3706 AGTTCTCTTAATGAGAAGTTGATTTTCGCAACCCACCAATATCACTTTCTTTGTTTAT 3765  
1257 ValSerAsnPheThrTTPProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276  
3766 GTCAGTAATTTCACTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCAGCAATTTT 3825  
1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuVal 1296  
3826 ATGACCTGTGTACAGTCTTTCGTGACTTCTTTCAGTGTGTTCTCTCTCTTCTGCTCTGGTG 3885  
1297 AlaAlaValValTTPLysIleLysGlnSerCysTTPAlaSerArgArgGluGlnLeu 1316

Db 3886 GTCTGCTGTTGGAGATCAACAAAGTGTGTGGCCCTCCAGAGCTAGAGCAACTT 3945  
 QY 1317 LeuArgGluWetGlnGlnMetA1aserArgPropheAlaserValAsnValAlaLeuGlu 1336  
 Db 3946 CTTTCAGAGATGCAACAGATGGCCAGCCGTCCTTTGCTCTGTAAATGTCGCCCTTGGAA 4005  
 QY 1337 ThrAspGluGluProProAspLeuLeuGlyGlySerIleLysThrValProLysProIle 1356  
 Db 4006 ACAGATGAGAGCGCTCCTCATCTTATTGGCGGAGTATAAGACTGTTCCCAAAACCCATT 4065  
 QY 1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376  
 Db 4066 GCAGTGGAGCGGTGTTTGGCAACAGCGCTGCTCTGTGTGTGGAGGCTCCCT 4125  
 QY 1377 ArgGlyLeuGlyGlyLeuProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396  
 Db 4126 CGAGGCGCTGGGTGATCCTCCTCTGGGAGTCAAGTCTTGTCTGGCCAGCGCCCTG 4185  
 QY 1397 ValAspIleSerGlnGlnValProIleValTyrLysGluLysSerGlyAlaValArgAsn 1416  
 Db 4186 GTGGACATTCTCAGCAGATGGCGTAGTGTACAAAGGAGAGTCAAGAGCGCTGAGAAAC 4245  
 QY 1417 ArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429  
 Db 4246 CGGAGCAGCAGCGCCCTGCACAGCTGGGACCTGCATC 4284

RESULT 2  
 AF5311101  
 LOCUS  
 DEFINITION Bos taurus attractin mRNA, complete cds.  
 ACCESSION AF5311101  
 VERSION AF5311101.1 GI:22297307  
 KEYWORDS  
 SOURCE Bos taurus (cow)  
 ORGANISM  
 Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.  
 1 (bases 1 to 4493)  
 Graphodatskaya, D., Joerg, H. and Stranzinger, G.  
 Direct Submission  
 Submitted (18-JUL-2002) Animal Sciences, ETH, Tannenstrasse 1,  
 Zurich, CH 8092, Switzerland  
 Location/Qualifiers  
 1..4493  
 /organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /chromosome="13"  
 52..4299  
 /note="transmembrane form; mahogany protein"  
 /codon\_start=1  
 /product="attractin"  
 /protein\_id="AA095445.1"  
 /db\_xref="GI:22297308"  
 /translation="VVAATAEARLRRTVTPAPTSRNGRRRCRDWDATRAGRPLRTGL  
 RLPRWRRLILLPLLLPAAEAAAAAASVSSAAAEAKECDPCVNGRCNPGL  
 GCVCPAGWQCHQCFRLTSGSGFTDGPNGYKTKTWLIEGQNPIMRLR  
 FNHFATCSWHLXYVDGSLYAPLVAAPFSLIVPERDNETPEVAVATGYALLHFF  
 SDAAYNLGPNITNFMPCNPNCSGRGECKTSNSNTVQCCESENWGEACDIPHCYN  
 NGCFPHRGICNSDVRGSCFSEWPGGCVFVPANQSFWRREYSDILKLPASHKAV  
 VNGNTMWWVGGMFNEHSDYNMVLAYDLASREWLALNRVSVVRYVGHSLALYQDKY  
 IFGHCPLGYLSVQVDELDRNTKTLHTQCALVQGGYVSDVDRTKALYIHGGYK  
 IFGHCPLGYLSVQVDELDRNTKTLHTQCALVQGGYVSDVDRTKALYIHGGYK  
 SFGANKYRLADLXYVDGSLYAPLVAAPFSLIVPERDNETPEVAVATGYALLHFF  
 SHGACFSSDFMAYDIDCDMSVLPRPLDHDVNRFGHSAVLNRTWVFGGNSLIL  
 SDILVFTSEQEAHSEARACLAGPGVRCMDTSSQCVSWELAPEAQEKIKSECFK  
 RIFDRDQDLTDCSTANTINGCQNDHCVPNTNHSCTEGQISIFKTDHCPKQNPY  
 YNKKTSRCSALDCNQCEPRNOECIALPENICIGIHWLVGNSCLKITARETYDNA  
 KLSICRNHAFIALSTTKKVFQKLRIQSSQMSKFTLTPWVGLRKINIVSWCYW  
 DMSPTNSLOWMSEPDAGFCGLISEPTSLGKAATCINPLNGSCERPANSAQ  
 CRTPCALATACECTSGSECMWCSNMKQCVDSNAYVASFPGOCMEWYTWSSCPEN  
 CSGYCTCSHCLEQPCGWCCTDPSNTGKCKIEGSKGPKVMPQSQPTGNSYPQLLNS

SMLEDSRYNWSFIHCPACQCNHSGKINQISICEKCNELTGHKCTCISGFYDPTN  
 GSKQCFRCNQHSLCNTNTKCFCTTKGVKDECOLCEVNRVQGNPKGTCYYILL  
 IDYQFTFSLSDRDRVYTAINFVATPDEQNRDLDMFINASKNFLNINIWAAGSFSAQTO  
 AGSEMPVSKTNIKEYKDSPEKFKFRNHNITFFVYVSNITWPIKIOIAFSOHNF  
 MDLYQFVFFFSFSLILLVAAVVKIKQSCWASRRERQQLREMOMQASRRPASVVA  
 LETDEPDDLIGSITKVPKPIALBPCFNGKAAVLSVFLPRGLGJLIPPOGSLAV  
 ASALVDISQMPVYVYKEKGAVERNKKQPPAQPGTCI"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 4493  
 Score: 7405.00 Matches: 1340  
 Percent Similarity: 96.35% Conservative: 33  
 Best Local Similarity: 94.04% Mismatches: 40  
 Query Match: 93.56% Indels: 12  
 DB: 4 Gaps: 5

US-09-787-097-12 (1-1429) x AF5311101 (1-4493)

QY 5 AlaAlaAlaThrGluAlaArgLeuA1ArgArgThrAlaAlaThrAlaAlaLeuAlaGly 24  
 Db 58 GCGGCGCGGCTGAGCGAGGCTG---AGGAGGACCGTGTGACACCGCGCCACGAGC 114  
 QY 25 ArgSerGlyGlyProHisTrpAspValThrArgAlaGlyArgProGlyLeuGly 44  
 Db 115 AGGACGCGAGCGGTGTGGGACTGGGACCGGACTAGAGCTGGAGAGCGCGGCTGCCG 174  
 QY 45 AlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArgLeuLeuLeu 64  
 Db 175 ACCGGGCTGGCTCCACGG-----CGGTGGCGGCGCTGCTGCTGCTG 216  
 QY 65 LeuLeuLeuLeuProProLeuLeuLeuLeuLeuLeuProCysGluAlaGluAlaAla 84  
 Db 217 CTGCTGCTC-----CGCGCG-----CTGCTGCTACTTCCCTGGCGGCTGAGGCGCG 264  
 QY 85 AlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgPro 104  
 Db 265 GCGGCGCGGCGTGGTGTGGGCTCAGCGCGAGCGAGGCAAGGAATGTGACCGGCC 324  
 QY 105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124  
 Db 325 TGTGTCAACGCGCGCGCTGCAACCCCTGGCACCGGCGAGTGTCTGCCCGCGCGCTCG 384  
 QY 125 ValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerGlyPhe 144  
 Db 385 GTGGCGGAGCAATGCCAGCAGCTGGCGGGCGCGCTTCAGACTAATGATCTTCTGATTT 444  
 QY 145 ValThrAspGlyProGlyAsnTyrLysThrLysThrLysCysThrTrpLeuLeuGly 164  
 Db 445 ATAACAGATGGACCTGGAATATAAATATAAATAAATAAATAAATAAATAAATAAATAA 504  
 QY 165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrLysCysSerTrpAsp 184  
 Db 505 CACCAAAATAGATAATAGACTTCCTTTCAATCAATTTGCTACAGAGTGTAGTTGGAC 564  
 QY 185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGly 204  
 Db 565 CATTTATATGTTTATGACGGGACTCAATTTATGACACCCCTAGTGTGCTGCTTACGGCG 624  
 QY 205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGly 224  
 Db 625 CTCAATGTTCTGAGAGACGCGCAATGAGACTGTCCCAGAGGTTGTCCCAATCAGGT 684  
 QY 225 TyrAlaLeuLeuHisPheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThr 244  
 Db 685 TATGCCCTGCTGCTATTTTATGATGCTGCTTATTAATCACTGACTGGAATTAATCACT 744  
 QY 245 TyrSerPheAspVetCysProAsnAsnCysSerGlyArgGlyCysLysIleSerAsn 264  
 Db 745 TACAATTTTGACATGTGTCCAAATTAATTTCTCAGCGCGAGGAGAAATGTAAGATCAGTAAT 804  
 QY 265 SerSerGluThrValGluCysGluCysSerGluAsnTyrLysGlyAlaCysAspIle 284

805 AGCAACACCTGTTCAATGTGAATGTTCTGAAAACTGAAAAGGTGAAGCATGTGACATT 864  
QY  
285 ProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVal 304  
Db  
865 CTTCACTGTGAATAACTGTGGTTTCTCTATCGAGGCATCTGCAATTCAGTGTATGTC 924  
QY  
305 ArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAla 324  
Db  
925 AGAGGATGCTCTGCTCTCTCAGATGGCAAGGTCTCGATGTTTCTGTGTACCAAGCT 984  
QY  
325 ArgGlnSerPheTrpThrArgGluCysThrAsnLeuValLeuProArgAlaSerHis 344  
Db  
985 AACCAAGTCAATTTGACACGAGAGGAATATCTGACCTAAAGCTCCCCAGAGCTCTCAC 1044  
QY  
345 LysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSer 364  
Db  
1045 AAAGCTGTGGTCAATGGAATATATATGTTGGTGTGTTGGCGGATATATGTTCAACCACTCA 1104  
QY  
365 AspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArg 384  
Db  
1105 GATTACAACTGGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTGCCTGAACCGT 1164  
QY  
385 SerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIle 404  
Db  
1165 TCTGTGAACAGTGTGGTTGTAGATATGGCCATCTTTAGCATATATACAGATAAAT 1224  
QY  
405 TyrMetTyrGlyLysIleAspProThrGlyAsnValThrAsnGluLeuArgValPhe 424  
Db  
1225 TACATGTATGGAGGAAAAATGATTCAACAGGGAATGTGACCAATGAGTTGAGAGTGT 1284  
QY  
425 HisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaVal 444  
Db  
1285 CATATTCAATGAATCATGGGTGTGTTGTCCTCCCAAGCAAGAGGAGCATATGACGTG 1344  
QY  
445 ValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIle 464  
Db  
1345 GTTGGGCATCGCGCACATTGTCACGTGTAATAATGGCCGAGTGGTCATGCTGTTATC 1404  
QY  
465 PheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGlyTyrAspLeuAspLys 484  
Db  
1405 TTGTGTCATGTCCTCTCTATGGATATATAAGCAGTGTGCAGGAATATGACTTGGATAAG 1464  
QY  
485 AsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSer 504  
Db  
1465 AACACATGGAGTATATTACACTCAGGGTGCCTTGTCCAAAGGGGGCTACGGCCATAGC 1524  
QY  
505 SerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSer 524  
Db  
1525 AGTGTCTATGATGATAGGACCAAGCCCTGTATATTCACGGAGGCTACAAGGCTTTTCACT 1584  
QY  
525 AlaAsnLysTyrArgLeuAlaAspLeuTyrArgTyrAspValAspThrClnMetTrp 544  
Db  
1585 GCCAATAAATACCGGCTTGCAGATGATCTCTACAGATATGATGTGGACCCAGATGGG 1644  
QY  
545 ThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGly 564  
Db  
1645 ACCATTCTTAAGACAGCCGATTTTCCGTACTTGCACACAGCTGTGATAGTGTGA 1704  
QY  
565 ThrMetLeuValPheGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLys 584  
Db  
1705 ACCATGCTGTTTGGAGGAACACACACAATGACAGCTCATGATGTGGACCCAGATGGG 1764  
QY  
585 CysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuPro 604  
Db  
1765 TGTCTCTCTCAGATTTATGCTTATGACATGATTGTGACCGATGTGTCAGTGTCTCC 1824  
QY  
605 ArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSer 624  
Db  
1825 AGACCTGATCTCCACCATGATGTCAACAGATTGGCCACTCAGCGGTCTTATACAACAGA 1884  
QY  
625 ThrMetTyrValPheGlyPheAsnSerIleLeuLeuSerAspIleLeuValPheThr 644  
Db  
1885 ACCATGATGATTTGGTGGTTTAAATAGTCTCCCTCCTCAGTGACATCTCTAGTATTACA 1944

QY  
645 SerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIle 664  
Db  
1945 TCGAAACAGTGTGAAGCACACACAGAGTGAAGCTGTGTCTCGCAGCAGGACCTGGTGT 2004  
QY  
665 ArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGlu 684  
Db  
2005 CGGTGTATGTGGACACAGGTCATCTCAGTGTGTCTCTCGGAGCTGGCA---CTTGAA 2061  
QY  
685 GlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCys 704  
Db  
2062 GCACAGAAAGATAAGTCAGATGTTTCTAAAGAATTTTTCACCATGACAGATGT 2121  
QY  
705 AspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsn 724  
Db  
2122 GACCAGCTCAGGATTTTACAGCTGCACAGCCACACCAATGGCTGCCAGTGGTCCAAT 2181  
QY  
725 AsnHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyr 744  
Db  
2182 GACATTTGTCTCCTACGAACACACAGCTGCACAGAAGGCCAGATCTCCATTTCAAGTAT 2241  
QY  
745 GluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSer 764  
Db  
2242 GATCATTGCCCCAAGGATAACCCGATGATTATTGTAAACAAGAAGACACAGCTCAGGAGC 2301  
QY  
765 CysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784  
Db  
2302 TGTGCCCTGACACAGAACTGCCAGTGGGAGCCCCGGAATCAGGAGTGCATCGCCCTGCC 2361  
QY  
785 GluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThr 804  
Db  
2362 GAAAAATCTGTGGCATTTGGCTGGCATTTGTTGGAACACTGTTTGAATAATTACTACT 2421  
QY  
805 AlalysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuAla 824  
Db  
2422 GCCAGGAGACTTATGCAATGCAAAATGTCTCGCAGAAACACCAATGCTTTTGGCT 2481  
QY  
825 SerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSer 844  
Db  
2482 TCTCTTCAACTCAGAAAGAGGTAGAAATTGTCTTAAAGCAGTGGCAATCATGCAGTCA 2541  
QY  
845 SerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnVal 864  
Db  
2542 TCACAGAGCATGTCCAAGTTCACTAAACCCCATGGCTTGGCTTCGGAAGATCAATGTG 2601  
QY  
865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884  
Db  
2602 TCTATTGGTGGTGGGAGATATGTCCCGTTCAAAATAGTTCACTTACAGTGGATGCCA 2661  
QY  
885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904  
Db  
2662 TCTGAGCCAGTGNATGCTGGATTCTGTGGGATCTTGTTCAGAACCCAGTACTCGGGGCTTA 2721  
QY  
905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHis 924  
Db  
2722 AAGSCTCAACCTGCATCAATCCACTCAATGGTAGTGTCTGTGAAAGGCCCTGCTAACCC 2781  
QY  
925 SerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSer 944  
Db  
2782 AGTCCCAAGCAGTGTGGACCCCTTGTCCCTTGGGACACCCCTGTGGGGAGTGTACAGT 2841  
QY  
945 GlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyr 964  
Db  
2842 GGCAGCTCGGAGTGATGTGGTGAGCAACATGAAGCAGTGGTGGACTCCAATGCTTAC 2901  
QY  
965 ValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProPro 984  
Db  
2902 GTGGCCTCCTTCCCTTTTGGCCAGTGTATGAGTGGTATACCATGAGCAGTGGCCCCCT 2961  
QY  
985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004  
Db  
2962 GAAAAATTGTCAGGCTACTGTACTGTAGTCACTGCTTGGGACCAACCAAGCTGTGGCTGG 3021

1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrIlysGlyPro 1024  
1024  
3022 TGCACTGATCTAGCAATACGGGCAAGGAAATGATGATGAGGGCTCTCTATAAGAGGACCA 3081  
1025 ValIysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044  
1044  
3082 GTGAAGATGCTTCTCAGGCGCTCAGGAAATCTCTACCCAGAGCTCTTCTCAATTC 3141  
1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGln 1064  
1064  
3142 AGCATGTGTCTGGAGACAGCAGATATACTGTGTCTTTCATTTCATGTCAGCTTGGCCAG 3201  
1065 CysAsnGlyHisSerIlysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThr 1084  
1084  
3202 TGAATGGCCACAGCAAAATGCAATATACAGATATCTGTGAGAGGTGTGAGAACCTGACC 3261  
1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGly 1104  
1104  
3262 ACGGCAAGCACTGTGAGACCTGCATATCTGGCTTCTATGTTGATCCCAACACGGGGA 3321  
1105 LysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLys 1124  
1124  
3322 AAATGTCAACCATGCAATGCAATGCAATGCTGTGATGCAATACCAACACGGGCAAG 3381  
1125 CysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsn 1144  
1144  
3382 TGCTTCTGCACCAACCAAGCGCTCAGGGGATGATGATGATGATGATGATGATGATGAT 3441  
1145 ArgTyrGlnGlnAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGln 1164  
1164  
3442 CGATACCAAGGAACCTCTCAAGGAACGCTGTATTATCTTCTTATTGACTATCAG 3501  
1165 PheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAla 1184  
1184  
3502 TTCACCTTTAGCCTATCTCCAGAGAGATGACCGCTATTACAGACCATCAATTTTGTGCG 3561  
1185 ThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerIysAsnPheAsn 1204  
1204  
3562 ACACCTGTATGACCAACCAAGGATTTGGACATGTTTCATCAATGCTTCCAAAACCTTCAAC 3621  
1205 LeuAsnIleThrTrpAlaIleSerPheSerAlaGlyThrGlnAlaGlyGluGluMetPro 1224  
1224  
3622 CTCAACATCACCTGGGCTGCCAGCTTCTCAGCTGGAACCCCAAGCTGGAGAGAGATGCT 3681  
1225 ValIleSerIysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAsp 1244  
1244  
3682 GTTGTTCACCAACCAACCAATTAAGGATGATCAAGATAGTTTCTTATGAGAGATTTGAT 3741  
1245 PheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIle 1264  
1264  
3742 TTTGCAACCAACCAACCAATATCTTCTTCTTGTATGATGATGATGATGATGATGATGAT 3801  
1265 LysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheVal 1284  
1284  
3802 AAAATTCAGATGATCTTCCAGCAGCAGCAATTTATGACCTAGTACAGATTCCTCGT 3861  
1285 ThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValTrpIlysIleLys 1304  
1304  
3862 ACCTTCTTCTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3921  
1305 GlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnValAla 1324  
1324  
3922 CAAAGTGTGTGGCTTCCAGAGCTGAGAGCAACTTCTTCCAGAGATGATGATGATGATGAT 3981  
1325 SerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeu 1344  
1344  
3982 AGTCGTCTCTTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 4041  
1345 IleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsn 1364  
1364  
4042 ATTGGGGGAGTATAAGACTGTCTCCAGAGCCCAATTCCTGAGCCGCTGTTTGGCAAC 4101  
1365 LysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProPro 1384

Db 4102 AAAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4161  
Qy 1385 ProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetPro 1404  
1404  
4162 CCTGGGAGTCCGGCTCGAGTGGCAGTCCCTGCTGGTGGACATTTCTCAGCAGATGCCA 4221  
Qy 1405 IleValTyrIlysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProAlaGln 1424  
1424  
4222 GTCGCTACAAAGAGAGAGTCCAGGAGCAGTGGAGAAACCGAAACAGCAGGCCCTGGCAG 4281  
Qy 1425 ProGlyThrCysIle 1429  
1429  
4282 CCTGGGACCTGCATT 4296  
RESULT 3  
AB062913 4284 bp mRNA linear ROD 16-JAN-2002  
LOCUS Mesocricetus auratus Atrn mRNA for attractin, complete cds.  
DEFINITION  
ACCESSION AB062913  
VERSION AB062913.1 GI:16930100  
KEYWORDS  
SOURCE Mesocricetus auratus (golden hamster)  
ORGANISM Mesocricetus auratus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
Mesocricetus.  
REFERENCE  
AUTHORS Kuramoto, T., Nomoto, T., Fujiwara, A., Mizutani, M., Sugimura, T. and Ushijima, T.  
TITLE Insertional mutation of the Attractin gene in the black tremor hamster  
JOURNAL Mamm. Genome 13 (1), 36-40 (2002)  
MEDLINE 21635551  
PUBMED 11773967  
REFERENCE  
AUTHORS Kuramoto, T. and Ushijima, T.  
TITLE Direct Submission  
JOURNAL Submitted (08-JUN-2001) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division, Tsukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail: tkuramoto@ncc.ncc.go.jp, URL: http://www.ncc.go.jp/research/rat-genome/, Tel: 81-3-3542-2511, Fax: 81-3-5565-1753)  
FEATURES  
Location/Qualifiers  
1..4284  
/organism="Mesocricetus auratus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10036"  
1..4284  
/gene="Atrn"  
1..4284  
/gene="Atrn"  
/codon\_start=1  
/product="attractin"  
/protein\_id="BAB72012.1"  
/db\_xref="GI:16930101"  
1  
/translation="MVAAAATALLRGSTAATAAPAGKRVQYRLCAATGAWRPGPR  
AGLCPLWVSRALPPPLLLLLLLPLPREAAAVSAASVGSAAAEACEDRP  
CVNGRCNPGTQCVCPAGVGBQCQCGRFLRTGSGFVTDGPNKYTKTWLI  
EGQPNIMRLRFNFHATECSWDHLYVVDGDSIVAPLIAAFSGLIIVPERDGNETAPEVT  
VTSQALLHFFSDFAAVNLATGNITYNFDMPCNCSGRGCKSSNSSTVECCSENWK  
GESCDIPHTDNCGPFHRGICNASDVRCSCSPHWGPGCSIPVPANQSFWRREYS  
KFLPRASHKAVVNGNIMWVGGMFNFHSDISWLAIDLASREWLPLNHYAVNVVRYG  
HSLAHKDKIYMYGKIDTGNVTNLFVHINESWLLTPAKEQYHVAQSHIV  
TLATRVWLVIPFGCHLYTISIVQYDLEKNTWSILQTOGALVGGYGHSSVTDHR  
TKALYVHGKAPFSAKRYLADLYVDVDTQVTLKQDSRFRYLVHTAVIVSGTMLV  
FGGNTDMSHGAKCFSSDFMAYDIACDRSVLPRDLHDVNRFPGHSAVLHNSSTM  
YFGTFNSLLISPLVTFEQDADHSEACVAAGPGICLWDVTSRRTCTWELATEE  
QABKLSGECFKTILDEDCDQDTCYCTANTNDCHNCNDHCPVNHSTCTSGOISL  
KYNDPCNPMTYCNKKTSCRCALQDNCQWEPNQCETALPENICGIIHVLGNSCL  
KITAKENDYNAKLSCRNHAFSLATSOKKVEFVLKHLRLMQSSQSMGKLLITPVG  
LRKINYSYMWEDMSPFTLQWMEPSDAGFCGILSEPTRGKAACTINPLNGS  
VCERPANHSKQCRTPCALRTACSECTSSSSSCSMWCSNMKQCVDSNAYVAFPPGQCM





1855 TTGCAACAAGTACCATGATGTTGCTGGGGTTTCAACAGCGCTCCTCCTCAGTGACGTC 1914  
QY  
641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAla 660  
DB  
1915 TTAGTCTTTTACCTCGGAGCAATGTGAGCCCAACCAAGTGGAGCGCTGTGTACAGCA 1974  
QY  
661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680  
DB  
1975 GGAACCTGGTATCGGTGTCTGTGGGACACAGTGTCTCATCTCGATGTACCTCTCGGAGTGT 2034  
QY  
681 AlaThrAspGluGlnGluLeuLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700  
DB  
2035 GCAACTGGAAGCAACAGCAGAAAAGTTAAATCAGAATGTTTTTCCAAAAGAACCCCTTGAC 2094  
QY  
701 HisAspArgCysAspGlnHisThrAspCysTrpSerCysThrAlaAsnThrAsnAspCys 720  
DB  
2095 CATGACAGATGTGACCAAGCACACAGATTTGTACAGCTGCACAGCCATACCAATGACTGC 2154  
QY  
721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
DB  
2155 CACTGGTCAATGATCACTGTGTCTGCTGTAACCAACAGCTGCACAGAGCCAGATCTCC 2214  
QY  
741 IlePheArgTrpGluAsnCysProLysAspAsnProMetTrpTrpCysAsnLysLysThr 760  
DB  
2215 ATTTCGAAGTATGCAATTCGCCCAAGGATAACCCCATGCTACTGCAATTAAGAAACA 2274  
QY  
761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnCys 780  
DB  
2275 AGCTGCAGAAAGCTGTGCTTAGACACAGAACTGCCAGTGGGAGCCCGGAATCAGAGTGC 2334  
QY  
781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800  
DB  
2335 ATCGCTCTGCCTGAAATATCTGTGCATTTGGTGGCATTTGGTGGAACTCATGTCTG 2394  
QY  
801 LysIleThrAlaLysGluAsnTrpAspAsnAlaLysLeuPheCysArgAsnHisAsn 820  
DB  
2395 AATACTACTCTGAAGAAATTAATGCAATGCTTAGTTCCTCTGTAGAACAACACAAAT 2454  
QY  
821 AlaLeuLeuAlaSerLeuThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840  
DB  
2455 GCCTTTTGGCTTCCCTTACATCCCAAGAAAGGTGAATTCGTCTTAAAGCATCTTGA 2514  
QY  
841 IleMetGlnSerSerGlnSerMetLysLysLeuThrLeuTrpTrpValGlyLeuArg 860  
DB  
2515 TTAATGAGTCACTCAGAGCATGCTCAAGCTCACTCTCACTCCATGGGTGGTCTTCGG 2574  
QY  
861 LysIleAsnValSerTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880  
DB  
2575 AAGATCAATGTGTCTTATTGGTCTGGAGGATATGCTCTCCATTTACAAATAGTTTGTGTG 2634  
QY  
881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
DB  
2635 CAGTGGATGCCATCCGAGCCAGTGTGCTGGTTCCTGTGGATCTTGTGAGGCTAGT 2694  
QY  
901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
DB  
2695 ACTCGGGGATTGAGGCTGCACCTGCATCAACCCACTCAATGGCAGTGTGTGTAAGG 2754  
QY  
921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
DB  
2755 CTGTCAAAACCAAGTGCAGAGTGCAGGACACCGGTGTGGCTCGGACAGCGTGTAGT 2814  
QY  
941 AspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960  
DB  
2815 GAGTGCATAGCAGAGCTCTGATGCTCATGTGGTGCAGCAACATGACAGTGTGTAGAC 2874  
QY  
961 SerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTrpTrpThrMetSer 980  
DB  
2875 TCCAATGCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAGTGTGTATCAATGAGC 2934  
QY  
981 ThrCysProGluAsnCysSerGlyTrpCysThrCysSerHisCysLeuGlnPro 1000  
DB  
2935 AGCTGCCACCTGAAATTTGCTGTGCTACTGTACTCTGAGCCATTTGCTTGGAGCAGCA 2994

1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020  
DB  
2995 GCGTGTGGCTGTGTACTGATCTTAGCATACTTGGGAAGGGAATGTATCATGGGGCAGC 3054  
QY  
1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTrpProGlnPro 1040  
DB  
3055 TATAAGGAGCCCATAGGATGCCTTACAGGCTTCTACAGAAATGTGTATCCACAGCCT 3114  
QY  
1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTrpAsnTrpSerPheIleHisCys 1060  
DB  
3115 CTTCTGAACCTCTAGCATGTGTTAGAGGACAGAGATACAACTGGCTTCTTCACTACTGT 3174  
QY  
1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGlyLysCys 1080  
DB  
3175 CCAGCTTGGCAGTCAATGACACAGCAATGATTAAACCAGAGTATCTGTGAGAAGTGT 3234  
QY  
1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTrpGlyAspPro 1100  
DB  
3235 GAGAACCTGACCTGCAAGCACTGTGACACTGCATATCTGCTTCTATGTGTATCCG 3294  
QY  
1101 ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120  
DB  
3295 ACTAATGAGGCAATGTGAGCCATGTAAGTGCACGGGATGATCGCTGTCAATACC 3354  
QY  
1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140  
DB  
3355 AACACCGGCAAGTGTCTGCACCAACAGGGTGTCAAGGGAGATGATGCCAGCTATGT 3414  
QY  
1141 GluValGluAsnArgTrpGlnGlyAsnProLeuArgGlyThrCysTrpTrpThrLeuLeu 1160  
DB  
3415 GAGTGAAGAAATCGATACCAAGGAACCCCTCTCAAGGAACATGTTATTATACCTTCTC 3474  
QY  
1161 IleAspTrpGlnPheThrPheSerLeuSerGlnGluAspArgTrpTrpThrAlaIle 1180  
DB  
3475 ATTGACTATCAGTTCACCTTTTAGCTGTCCAGGAAGACGATCGCTATTATACAGGCATC 3534  
QY  
1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200  
DB  
3535 AATTTTGTGGCCACTCTGTATGAACAAACAGGGATTTGGACATGTTTATCAATGCTCA 3594  
QY  
1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220  
DB  
3595 AAAAACTTCAACCTCAACATCACCTGGCTTACAGCTTCTCAGCTGGAACCCAGGACGA 3654  
QY  
1221 GluGluMetProValValSerLysThrAsnIleLysGluTrpLysAspSerPheSerAsn 1240  
DB  
3655 GAAGAAATGCTGTGTGTTTCAAAAACCAACATTAAGAGGTACAAAGATAGCTTCTCTAAT 3714  
QY  
1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTrpValSerAsnPhe 1260  
DB  
3715 GAGAAATTTGATTTTGGCAACCATCCAAACATTTCTTCTTCTTCTTATGTCACTAATTC 3774  
QY  
1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280  
DB  
3775 ACCTGGCCCATCAAAATTCAGATTGCTTCTCCAGCACAGCAATTTATGGACCTGGTA 3834  
QY  
1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValVal 1300  
DB  
3835 CAGTTCTCTGTGACTTCTTCTCAGCTGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3894  
QY  
1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMet 1320  
DB  
3895 TGAAGATCAAGCAGAGCTGTGGCATCCAGGAGGAGGAGCAACTTCTTCTGGGAGATG 3954  
QY  
1321 GlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340  
DB  
3955 CAGCAGATGCCAGCCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4014  
QY  
1341 ProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360  
DB  
4015 CCTCTCATCTTATTGGGGGAAGTATAAAGACTGTCCCAAGCCCATTTGCCCTGGAGCCC 4074



```
Qy 1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380
Db 4075 TGCTTTGGTAAACAAGCGCGAGTCTCTCTGTTATTTGAGCGCTCCCTCGAGGACTGGGT 4134
Qy 1381 GlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer 1400
Db 4135 GGAATCCCTCCTCTGCTGCTCAGTCTGCTGTGGCCAGTGGCTGCTGCTGCTGCTGCT 4194
Qy 1401 GlnGlnMetProIleValTyrIleGlySerGlyAlaValAlaArgAsnArgGlyGlnGln 1420
Db 4195 CAACAGATGCCGATAGTATACAAAGAGAGTCAAGAGCGGTAAAGAAACCGGAAGCAGCG 4254
Qy 1421 ProProAlaGlnProGlyThrCysIle 1429
Db 4255 CCCCTGACACAGCTCGAACCCTGCAAT 4281

RESULT 4
AF116897 6083 bp mRNA linear ROD 20-MAR-1999
LOCUS Mus musculus mahogany protein mRNA, complete cds.
DEFINITION AF116897
ACCESSION AF116897
VERSION AF116897.1 GI:4454560
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Nagle, B.L., McGrail, S.H., Vitale, J., Woolf, E.A., Dussault, B.J. Jr.,
DiRocco, L., Holmgren, L., Montagnò, J., Bork, P., Huszar, D.,
Fairchild-Huntress, V., Ge, P., Keilty, J., Ebeling, C., Baldini, L.,
Gilchrist, J., Burn, P., Carlson, G.A. and Moore, K.J.
The mahogany protein is a receptor involved in suppression of
obesity
Nature 398 (6723), 148-152 (1999)
JOURNAL Nature 398 (6723), 148-152 (1999)
MEDLINE 99184159
PUBMED 10086355
REFERENCE 2 (bases 1 to 6083)
AUTHORS Nagle, D.L., McGrail, S.H., Vitale, J., Woolf, E.A., Dussault, B.J. Jr.,
DiRocco, L., Holmgren, L., Montagnò, J., Bork, P., Huszar, D.,
Fairchild-Huntress, V., Ge, P., Keilty, J., Ebeling, C., Baldini, L.,
Gilchrist, J., Burn, P., Carlson, G.A. and Moore, K.J.
Direct Submission
Submitted (28-DEC-1998) Sequence Analysis, Millennium
Pharmaceuticals, 640 Memorial Drive, Cambridge, MA 02139, USA
FEATURES
Location/Qualifiers
1..6083
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
70..4488
/note="MG"
/codon_start=1
/product="mahogany protein"
/protein_id="A4D20947.1"
/db_xref="GI:4454561"
/translation="MVAVAAATAEARLGRSTTTTAAAGRKGRHRRCTATGAWRPG
PRARLCPEVRALPPPLLEPLLELLPLPREABAAVAAVAGSAAAEKCDR
PCVNGRCNPGTGCVCPTGWGEQCHGGRFRTGSGFVDPGPNYKTKTTL
IEQPNRLMRFNFHAFECSDMDLYVDGSIYAPLIAAFSLIVPERDNETAPEV
TVTSVGLAHFFSDAAYNLTFNITNFDMPNCCGRGECKSSNSAVECESENW
KGESCDI PHTCNCGPFHRCICNASDTGSCFCFPHWQGGSIIPVPAQNSFWTREYS
DGLPASHAKVNGNIMWVGGYNFHSDYSNWLAYDLTSREWLPLNHSVNSVVRV
GSLALHDKDIYMGKIDSTGNINELRPHFHNSWLLTPKAKQIATVVGSAHI
VLAGRVVMVI FGHCPLYGISVQQEYDLEKMTWSILHTQALVQGGYGHSSYDD
RTKALYVHGCVKAFKANKYRLADDLYVDVTQMTILKDSRFFRYLHTAVTSGTML
VFGNTHNDSHGAKCFSSDMAIDACDWSVLRPELHVDNRFCHSAVLYNST
MVVFGFNSLLISDLVFTSEOCDAHREAAVAAAGPGIRCLMDTOSRCKSWELATE
EQAELKSCSPKRLDHRDCHQDTCYCTANTNDCHWCDHCPVNHSCTEGQISI
AYTESCPKDNPTMYCNKTSRCSALDQNCQWEPNQECIALPENICNGHLYGNCS
LKITTAKENYDNALSKNSNAFLASLTQKQBEVLKQLRMQSSSKSLTLPWV
GLRKTINVSCHWEDSPFTNSLLOWPSEPSDAGFCGLISEPSTGLKAATCINPLNG
SVCERPANHSKQRCPTPCALRTACGECTSSSECEMWCNSNMKQCVDSNAYVAFSPFGC
```

## ORIGIN

```
Alignment Scores:
Pred. No.: 0 Length: 6083
Score: 7346.50 Matches: 1332
Percent Similarity: 95.12% Conservative: 32
Best Local Similarity: 92.89% Mismatches: 56
Query Match: 92.82% Indels: 14
DB: 10 Gaps: 3

US-09-787-097-12 (1-1429) x AF116897 (1-6083)
Qy 2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAlaAla 21
Db 79 GTGGCGGCGCGCGCGCGACTGAGCGCGCGTGGGGGAGCAGCAGCAGCAGCAGCG 138
Qy 22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpValThrArgAlaGlyArgPro 41
Db 139 CCTGGCGGCGAGGAAGGCGAGCAGCAGCCCTGCACCGCAGCGGCGCTGGAGCGCG 198
Qy 42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuSer-----ProPro 56
Db 199 GGAACCGCGCGCGCGTGTCTCCCGCGGCGTGTGGCGGCGCGTGCCTGCGCGCGCG 258
Qy 57 LeuArgProArgLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
Db 259 CTGCTGCGCG---CTGCTCTTTTCGCTGCTGCTGCGCG-----294
Qy 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
Db 295 CTGCCCCGCGAGCGCGCGCGCGTGGCGCGCGCGCGTGCCTGCGCGCGCGAGCC 354
Qy 97 GluAlaLysGluCysAspArgProCysValAsnGlyValArgCysAsnProGlyThrGly 116
Db 355 GAGGCCAAGGAATGTGACCGCGCGTGTGTCAACGGCGCGCGCTGCAACCTCGCACCGGC 414
Qy 117 GlnCysValCysProAlaGlyTrpValGlyGlnCysGlnHisCysGlyGlyArgPhe 136
Db 415 CAGTGCCTGCGCGCGCGCGTGTGTCAACGGCGCGCGCTGCAACCTCGCACCGGCCTTC 474
Qy 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrThr 156
Db 475 AGACTAACTGCGCTCTTCTGGAATTTGTAAACAGATGGACCTGGGAATTAATAATTAAGACG 534
Qy 157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db 535 AAGTGCACATGGCTCATTTGAAGGACAGCCAAATAGATAATAGACTTCGCTTCAACCAT 594
Qy 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
Db 595 TTTGCTCAGAAATGTAGCTGGACCAATTTATATATATATATATATATATATATAGCGCA 654
Qy 197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
Db 655 CCTCTGATTGCTGCTTTAGTGGCTCATTTGCTTCTGAAAGAGATGGCAATGAGACGGCT 714
Qy 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236
Db 715 CCTGAGGTCACTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTTAT 774
Qy 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
Db 775 AATCTGACTGATTTATATATATATATATATATATATATATATATATATATATATATATCTCAGCG 834
```

QY 257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276  
DB 835 CGAGGAGAGTGTAAAGACAGTAAACAGCAGCGCTGTTGAGTGAATGTTCTGAAAC 894  
QY 277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296  
DB 895 TGAAGAGGGAGTGTGTGACATTCCTCACTGTACAGACAACTGTGGCTTCTCACCAG 954  
QY 297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316  
DB 955 GGCATCTGTAATCAAGACGATACAGAGGCTGCTCTCTCTTCTCCTCAGTGGCAGGCTCT 1014  
QY 317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGlyTrpSerAsn 336  
DB 1015 GGATGTTCAATTCCTGTGCGACGTAACCACTCTTTTGGACTCGAGAGAAATTTCTGAT 1074  
QY 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356  
DB 1075 TTAAGCTTCCAGAGCCTCTCATAAAGCTGTGGTCAATGGAAATATATATGTGGTGT 1134  
QY 357 GlyGlyTrpMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376  
DB 1135 GCGGATATATGTTCAACCACTTACAGATACAGATGCTTTAGCGTATGACTGACTTCT 1194  
QY 377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValArgTyrGlyHisSer 396  
DB 1195 AGGGAATGGCTCCCACTAAACCACTCTGTGAACAGTGTGGTGTGAAGATATGCTCAT 1254  
QY 397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416  
DB 1255 TTGGCATTACATAGGATAAATCTACATGTATGAGGAAATTTGATTCACAGGGAAC 1314  
QY 417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436  
DB 1315 GTACCAATGAGTGAAGATATTTCAATATTCATATGAATCAATGAGTATTTAACTCCG 1374  
QY 437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456  
DB 1375 AAAGCTTAAGATACAGTATGATGAGTGTGGACACTCAGACACATTTGTTACACTGGCA 1434  
QY 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476  
DB 1435 GCGCGTGTGTGTCATGTTGTCATCTCGGTCAATGCCCACTATGAGATATATAAGCGTT 1494  
QY 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496  
DB 1495 GTGCAAGGAATATGACTTGGAAAGAACACATGGAGTATATATACATPACTCAGGGTGTCT 1554  
QY 497 ValGlnGlyTyrGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516  
DB 1555 GTGCAAGGGGTTATGGCCACAGTAGTGTATATGATGACAGACCAAGGCTCTGTACGTT 1614  
QY 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536  
DB 1615 CATGTTGGCTACAGGCTTTACAGCGCCCAACAAATACCGGCTTGACAGATGACCTCTACAG 1674  
QY 537 TyrAspValAspThrGlnMetTrpThrIleLysAspSerArgPheArgTyrLeu 556  
DB 1675 TACGATGTGGTACTCAGATGTGGCAATCTTAAAGGACAGCGCAATTTTCCGTTACTTGT 1734  
QY 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576  
DB 1735 CATACAGCTGTATAGTGAAGTGAACCACTGTGTTGTTGGAGGAAACACACACAAATGAC 1794  
QY 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596  
DB 1795 ACTTCATGAGCCAGCGTGGCAATGCTTCTCCCTCGGACTTCATGGCTTATGACATGCT 1854  
QY 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616  
DB 1855 TGTGACCGATGTCAGTGTCTCCAGACCTGAGCTCCATCATGATGATCAACAGATTTGGC 1914  
QY 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636

DB 1915 CATTACAGCACTTGTACACACAGCAATGATGTGTGGCGGCTTCAACAGCCTCCTC 1974  
QY 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
DB 1975 CTCAGTGACCTTGTGTTTACCTCGGACAGTGCATGCACACGCGAGTGAAGTGTCT 2034  
QY 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676  
DB 2035 TGTGTGCACAGACACTGTGTATCCGCTGTCTGGGACACACAGTCTGCTCATGATACC 2094  
QY 677 SerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696  
DB 2095 TCCTGGGAGTGGCACTGAAGAAACAGCAGAAAGTAAATACAGAGTGTCTTCTAAA 2154  
QY 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
DB 2155 AGAACCTTGCACATGACAGATGTACACACACAGATTTGTTACAGCTGCACAGCAAT 2214  
QY 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
DB 2215 ACCAATGACTGCCACTGTGCAATGATCACTGTGCTCCTGTGAACACACAGCTGCAGAA 2274  
QY 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
DB 2275 GGCAGATCTCCATTTGCCAAGTATGAGAGTGTGCCCAAGGATAAACCCTACTACTGC 2334  
QY 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776  
DB 2335 AATAAGAAACACAGCTGCAGAGCTGTGCCCTAGACACAGACTGCCAGTGGGAGCCCCG 2394  
QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796  
DB 2395 AATCAAGAGTGCATGCCCTGCCGAAATATCTGTGGCAATGGCTGGCATTTGGTTGA 2454  
QY 797 AsnSerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
DB 2455 AACTCGTGTCTGAAATCACTACTGTGAAGAGAAATATGACAATGCTAAATTTGCTCT 2514  
QY 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836  
DB 2515 AGCAACCAATGCCCTTTTGGCTTCCCTCACATCCACAGAGAGGTGGAGTTGTCCTT 2574  
QY 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856  
DB 2575 AAGCAGCTTCGATTAATGAATCATCTCAAGTATGTCCAAGTCACTCTGACTCCATGG 2634  
QY 857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876  
DB 2635 GTTGTCTTCGAGNAGATCAATGTGCTTACTGTGTGGAGAGATATGCTCTCATTTACA 2694  
QY 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
DB 2695 AATAGTTTGTGTCAGTGGATGCCATCTGAGCCCACTGATGTGGCTTCTGTGGGATCTTG 2754  
QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
DB 2755 TCAGACCTTAGTACTCGGGATTAAGGCTGCAACTGATCAACCTCTCATGGCAGC 2814  
QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
DB 2815 GTCTGTGAAGGCTGCAAAACACACAGTGCAGAGAGTGCAGACACCACTGTCCCTGGG 2874  
QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956  
DB 2875 ACAGCTGTGGCAGGACAGTACAGCAGCTCGAGTGCATGTGTGTCAGTAACTGAAG 2934  
QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrp 976  
DB 2935 CAGTGTGTGAGTCCCAATGCCCTACGTGGCTCTCTCCCTTTTGGCCAGTGTATGAATGG 2994  
QY 977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996



GRSLKLTAKENYDNKLSCRNHNAPLASLTSQKVPFLVKQLRLMOSSQSTSLTL  
 TPVGLRKINVSINWCHWEDSPFINSJLQMPSPDAGFCGLSSEPSRTRGUKAATCIN  
 PLNGSCERPAHSAKQCTPCALRTACCECTSSSEBCWCSNMKQCYDSNAYVASFP  
 FOCMEWYTMSPCCPENCSCYCTCHLEQPCGWCCTDFSNITGKQKCIEGSYKGPVM  
 PSHASTNTYPPOLPLNSMCLDSRYNWSFIHCPACQCNHSCINQSIKERCEDLTT  
 GKHCETCISGFYDPTNGCKQPCCKNGHASLNTYNTGKFCCTKGVAGECOLCEVE  
 NRVOGNPLAGTCTYLLIDYOFTELSQEDDRYTAINFVAPEDQNDLDMFLNASK  
 NENLNTWATSPPAGTCTCEEPVVSNTKNIKEYKDSFNSKEDFRNHNPIFFVYVSN  
 FTWPIKIQAFSPGSMFLDVGFTVFTFSCFLSLLVAAVWVKIKQCSWASRRRLQL  
 REMQNASRPPASVNVALSTDEBPDLIGISIKTPKPIALEPFCFNGKAAVLSVEPLL  
 PRGLGIPPPGSGSLAVASALVDIISQMPDIVYKESGAVRNKQPPAPGCTCI"

## ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 8739  
 Score: 7342.50 Matches: 1331  
 Percent Similarity: 95.25% Conservative: 33  
 Best Local Similarity: 92.95% Mismatches: 65  
 Query Match: 92.77% Indels: 3  
 DB: 10 Gaps: 1

US-09-787-097-12 (1-1429) x AB038387 (1-8739)

QY 1 MetValAlaAlaAla-----AlaThrGluAlaArgLeuArgArgThrAla 17  
 Db 61 ATGGTGGCGCGCGCGCGCGCGAGGCGACTGAGGCGCGGTTGAGGGGTACACGAGC 120  
 QY 18 AlaThrAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 37  
 Db 121 GCACAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 180  
 QY 38 AlaGlyArgProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProLeu 57  
 Db 181 GCGTGGAGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 240  
 QY 58 ArgProArgLeuLeuLeuLeuLeuLeuLeuLeuLeuProProLeuLeuLeuLeuLeu 77  
 Db 241 TCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 300  
 QY 78 ProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 97  
 Db 301 CCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360  
 QY 98 AlaLysGluCysAspArgProCysValAlaAlaGlyArgCysAsnProGlyThrGlyGln 117  
 Db 361 GCCAAGGAATGTGACCGCGCGTGTCTCAACGCGCGCGCGCGCGCGCGCGCGCGCG 420  
 QY 118 CysValCysProAlaGlyTyrValGlyGluGlnCysGlnHisCysGlyGlyArgPheArg 137  
 Db 421 TCGGTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 480  
 QY 138 LeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysThrLys 157  
 Db 481 CTAACTGGATCTCTCGAATTTGTAACAGATGACCTGGGAATTAATAATAAACAAGAA 540  
 QY 158 CysThrTrpLeuLeuGluGlnProAsnArgIleMetArgLeuArgPheAsnHisPhe 177  
 Db 541 TCACATGGCTCATTGAAGGACGCGCAATGAATGAATGAATGAATGAATGAATGAATGA 600  
 QY 178 AlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaPro 197  
 Db 601 GCTACTGAATGTAGCTGGGACCATTTATATGTTTATGATGGGACTCAATTTACGCACCT 660  
 QY 198 LeuValAlaAlaPheSerGlyLeuLeuValProGluArgAspGlyAsnGluThrValPro 217  
 Db 661 CTGATTGTGCTTTAGTGGCTCTATTGCTGAAGAGATGGCAATGACAGCGGCCCT 720  
 QY 218 GluValValAlaThrSerGlyTyrAlaAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsn 237  
 Db 721 GAAGTCACTGTCACCTTCAGTTATGCACTACTGCAATTTTTCAGTGATGCTGCTTAAT 780  
 QY 238 LeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArg 257

Db 781 CTGACTGGATTTAATATCACTTACAATTTTGACATGTGCCGAATTAATTTCTCAGGCCGA 840  
 QY 258 GlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrp 277  
 Db 841 GGAGATGTGAAGACAGCACTAAACAGCAGCAGCACTGTTGAGTGTGAATGTTCTGAAAACTGG 900  
 QY 278 LysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGly 297  
 Db 901 AAAGGGGAGTCTGTGATCTCTCACTGTACAGACAACTGTGGCTTCTCTCACCAGGC 960  
 QY 298 IleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGly 317  
 Db 961 ATCTGTAAACGAAGTGACACACAGAGGGTGTCTCTGCTTCCCTCAGTGGCAGGGTCTCTGGA 1020  
 QY 318 CysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeu 337  
 Db 1021 TGTTCATTTCTGTGCCAGCTAAACAGTCTTTTGGAGCTCGAGAAGATATATCTGACTTA 1080  
 QY 338 LysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGly 357  
 Db 1081 AAGCTTCCAGAGCCTCTCATAAAGCTGTGCTCATGGAAATATATGTGGTGTGTGGC 1140  
 QY 358 GlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArg 377  
 Db 1141 GGGTATATGTTCAACCATTCAGATTAACAGTGTGTTAGCGTATGACCTGGCTTCTAG 1200  
 QY 378 GluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeu 397  
 Db 1201 GAATGGCTTTCATTAACCATTCGTGAACAGTGTGTTGTAAGATATGTCATTCTTTG 1260  
 QY 398 AlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnVal 417  
 Db 1261 GCATTACATAAGGATAAATTCATCATGTATGGAGGAAAAATTCATTCAACAGGAAATGTG 1320  
 QY 418 ThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLys 437  
 Db 1321 ACCAAGCAGCTGAGAGTATTTCATATATTAATGAATCATGGGTATTGTTAAACCCCAAA 1380  
 QY 438 AlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGly 457  
 Db 1381 GCTAAGGATCAGTATGACGTGGTGGACACTCAGCACACATTTGTACACTGTCTATGCG 1440  
 QY 458 ArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrLysSerAsnVal 477  
 Db 1441 CGTGTGCTCATGTGTGTCATCTTCGTCATCTGTCGCTCTACGGATATATAAGTGTGTG 1500  
 QY 478 GlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuVal 497  
 Db 1501 CAGGAATATGACTTGGAAAGAACACATGAGTATATTAACAACCTCAAGGTCTCTAGTG 1560  
 QY 498 GlnGlyGlyTyrGlyHisSerValTyrAspHisArgThrArgAlaLeuTyrValHis 517  
 Db 1561 CAAGGGGGTATGGCCACAGCAGTGTATTATGATCAGACCAAGGCTCTGTATGTTCAT 1620  
 QY 518 GlyGlyTyrLysAlaPheSerAlaLeuLysTyrArgLeuAlaAspLeuTyrArgTyr 537  
 Db 1621 GGTGGCTACAGGCTTTTCAGCGCCCAAGTACCGCTTGGCAGATGACCTCTACAGATAC 1680  
 QY 538 AspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHis 557  
 Db 1681 CATGTGGACACTCAGATGTGGACCATTCCTTAAGGACAGCGGATTTTTCGTTTACITGCAT 1740  
 QY 558 ThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThr 577  
 Db 1741 ACAGCTGTGATAGTGGAGCAACCATGCTGGTGTTCGGAGGGAACACACACAAATGACACT 1800  
 QY 578 SerMetSerHisGlyAlaLysCysPheSerAspPheMetAlaTyrAspIleAlaCys 597  
 Db 1801 TCCATGAGCCAGGTGCCAAATGCTCTCTCAGACTTCATGCTTATGACATGCTTGT 1860  
 QY 598 AspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHis 617  
 Db 1861 GACCGATGTGCTAGTTCTCCCGAGCCTGAGCTCCATCAGATGTCAACCGATTTGCGCCAC 1920

QY 618 SerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeu 637  
Db 1921 TCAGCAGCTTTCGCAACAGCAGCAGATGATGTTGTCGGCGCTTCAACAGCCTCTCTCTC 1980  
QY 638 SerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCys 657  
Db 1981 AGTGACGCTTGTAGTCTTCACTTCGAGCAGAGTGTGATGACACACCGAGCAAGCTGCTTGT 2040  
QY 658 LeuAlaAlaGlyProGlyIleArgCysValTriaAsnThrGlySerSerGlnCysIleSer 677  
Db 2041 GTGGCAGCAGACCTGGCATCGGTGTCTGTGGGACACACAGTGTCTCGATGTACCTCC 2100  
QY 678 TrpAlaLeuAlaThrAspGluGlnGluGlyLeuLeuLeuSerGluCysPheSerLysArg 697  
Db 2101 TGGGAGTGTGGCACTGGAAGACAGCAGAAAGTTAAATCAGAAATGTTCTCTAAAGA 2160  
QY 698 ThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThr 717  
Db 2161 ACCCTTGACCATGACAGATGTGACCAGACACAGATTTGTTACAGCTGCACAGCCAATACC 2220  
QY 718 AsnAspCysHisTyrCysAsnAspHisCysValProArgAsnHisSerCysSerGluGly 737  
Db 2221 AATGACTGCACCTGTGCAATGATCACTGTGTCTGTGAAACACAGCTGCACAGCAAGGC 2280  
QY 738 GlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsn 757  
Db 2281 CAGATCTCATTTGCCAAGTATGCAATTTGCCCAAGGATTAACCCCATGTACTACTGCAAT 2340  
QY 758 LysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTyrGluProArgAsn 777  
Db 2341 AAGAAACCCAGCTGCAGAGCTGTGCCCTAGATCAGAAATTCAGCTGGGAACCTCGAAAT 2400  
QY 778 GlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTyrHisLeuValGlyAsn 797  
Db 2401 CAAGAGTGCATCGCCCTGCCGAAACATCTGTGCATTTGGCTGTGGATTTGTTCGGAAC 2460  
QY 798 SerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArg 817  
Db 2461 TCGTGTCTGAATCACTACTCTAAGGAAATTTATGACATGTCTAAATTTGCTGTAGA 2520  
QY 818 AsnHisAsnAlaLeuLeuAlaSerLeuThrGlnLysLysValGluPheValLeuLys 837  
Db 2521 AACCAACATGCTTTTGGCTTTCGCTTACGTCCAGAAAGAGTGGAGTTGTCTCTAAG 2580  
QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTyrVal 857  
Db 2581 CAGCTTCGATTAATGCAATCACTCAGACACAGTCCAGCTCACTCTGACTCCATCGGTT 2640  
QY 858 GlyLeuArgLysIleAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsn 877  
Db 2641 GGTTCTCGAAGATCAATGATCTTACTGTGTGGAGGATATGTCTCATTTCAAAAT 2700  
QY 878 SerLeuLeuGlnTyrMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
Db 2701 AGTTCTGCTCAGTGGATGCCATCTGAGCCAGTGTGCGCGCTTCTGTGGATCTTGCA 2760  
QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917  
Db 2761 GAGCTTAGTACTCGGGAGCTGAAGCTGCAACCTGCATCAACCCACTCAATGGAAGTGT 2820  
QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937  
Db 2821 TGTGAAGGCTGCACAAACACAGTGCAGAGTGCAGACACCTGTGCCCTGCGGACA 2880  
QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLysGln 957  
Db 2881 GCATGTGGTGGTGCACACAGCAGCTCCGATGATGTGGTGACAGACATGAGCAG 2940  
QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyr 977  
Db 2941 TGGGTAGACTCCAAATGCTACGTAGCTCTCTTTCTTTCGGCCAGTGTATGGAGTGTAT 3000

QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
Db 3001 ACATGAGCAGCTGCCCACCTGAAATTTGCTCTGGCTACTGTACCTGCAGCCATTTGCTTG 3060  
QY 998 GluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017  
Db 3061 GAGCAGCCAGGCTGTGGTGGTGTACTGATCCTAGCAATACTGGGAAGCAAAATGATTT 3120  
QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
Db 3121 GAGGGAAGCTATAAGAGCCGCGGAAGATGCCCTTTCATGCCCTCTACAGGAATGTGTAC 3180  
QY 1038 ProGlnProLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTyrSerPhe 1057  
Db 3181 CCACAGCCCTTCTGAACCTCCAGATGTGTGTAGAGACAGAGATACAACTGGTCTTTT 3240  
QY 1058 IleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCys 1077  
Db 3241 ATTCACTGTCCAGCTTGCAGTGCATGTGACACAGCAAGTGCATCAACAGACATCTGT 3300  
QY 1078 GluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyr 1097  
Db 3301 CAGAGTGTGAGACCTGACCAAGCAAGCACTGCGAGACCTGCAATATCTGCTCTTAT 3360  
QY 1098 GlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117  
Db 3361 GGTGACCCGACTAATGGAGGCAATGTGAGCCATGCAAGTGCACCGGCGACGATCACTG 3420  
QY 1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137  
Db 3421 TGCACACCAACACTGGCAAGTGTCTGTACCAACAGGCGCTCAAGGAGGAGGTGC 3480  
QY 1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157  
Db 3481 CAGCTATGTGAGGTAGAAAAATCGATACACAGGAAACCTCTCAAGGAAACATGTTACTAT 3540  
QY 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyr 1177  
Db 3541 ACCCTCTCATTTGACTATCAGTTCACTTTTAGCTGTCCAGGAAGACAGCCGCTATTAC 3600  
QY 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197  
Db 3601 ACAGCCATCAATTTTGTGGCTACTCTGTATGAACAAACAGGATTTGGACATGTTTCATC 3660  
QY 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTyrAlaAlaSerPheSerAlaGlyThr 1217  
Db 3661 AATGCTTCCAAAACTTCAACCTCAACATCACTGGGCCCACAGTTTCCCGCTGGAAAC 3720  
QY 1218 GlnAlaGlyGluGluMetProValLysSerLysThrAsnIleLysGluTyrLysAspSer 1237  
Db 3721 CAGACTGGAGAGAGGTGCTGTGTGTTTCAAAAACCAACATCAAGGATACAAAGATAGC 3780  
QY 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257  
Db 3781 TTCTTAATGAGAAATTTGATTTTGCACCATCCAAACATTTACTTTCTTTGTTTATGTC 3840  
QY 1258 SerAsnPheThrTyrProIleLysIleGluIleAlaPheSerGlnHisSerAsnPheMet 1277  
Db 3841 AGTAATTTCACTTGGCCCATCAAAATTCAAATTCCTTCTCCACAGCAGCAATTTATG 3900  
QY 1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAla 1297  
Db 3901 GACCTGTGTGAGTCTTCGTGACCTTTCTCAGCTGTTTCTCTCTCTCTCTCTCTCTCTCT 3960  
QY 1298 AlaValValTyrLysIleLysGlnSerCysTyrAlaSerArgArgArgGluGlnLeuLeu 1317  
Db 3961 GCAGTGTGTGGAGATCAAGCAGAGCTGTGGGCAATCCAGAGGAGAGACAACTTCTT 4020  
QY 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337  
Db 4021 CGGAGATGCAACAGATGGCCGCGCTTTGCTTCTGTAAATGTTGCTTCTGGAACA 4080  
QY 1338 AspGluGluProProAspLeuIleGlySerIleLysThrValProLysProIleAla 1357

```

Db      4081 GATGAGAACCTCTGATCTCATTTGGGGAGATATAAGACTGTTCTTAAGCCCATTC 4140
Qy      1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377
Db      4141 CTGGAGCCCTGCTTGTGTAACAAGCTGCTCTCTCTGTGTGTTGTGAGGCTCCCTCGA 4200
Qy      1378 GlyLeuGlyClyLeuProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397
Db      4201 GGACTGGAGGAATCCCTCTCTCTGTCAGTACAGTCTGCTGTCGTCAGTCCCTGGTG 4260
Qy      1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417
Db      4261 GACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAGTCAGAGCTGTAGGAACCGG 4320
Qy      1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db      4321 AAGCAGACGCCCTCCACAGCTGGAACCTGCATT 4356

RESULT 6
AR164807
LOCUS      AR164807               8827 bp      DNA      linear      PAT 17-OCT-2001
DEFINITION Sequence 1 from patent US 6274339.
ACCESSION  AR164807
VERSION     AR164807.1  GI:16238010
KEYWORDS
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   Unclassified.
AUTHORS     Moore,K. and Magle,D.Lynn.
TITLE       Methods and compositions for the diagnosis and treatment of body
            weight disorders, including obesity
JOURNAL     Patent: US 6274339-A 1 14-AUG-2001;
FEATURES    Location/Qualifiers
            source          1..8827
                        /organism="unknown"
                        /mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      8827
Score:           7303.00      Matches:    1332
Percent Similarity: 90.99%      Conservative: 32
Best Local Similarity: 88.86%      Mismatches: 56
Query Match:      92.27%      Indels:    79
DB:               6      Gaps:      4

US-09-787-097-12 (1-1429) x AR164807 (1-8827)
Qy      2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAlaAla 21
Db      83 GTGGCGCGCGCGCGCGCGCTGAGCGCGGCTGAGCGGGGAGCAGCAGGACGACGACG 142
Qy      22 LeuAlaGlyArgSerGlyGlyProHisTrpAspValTrpAspValTrpArgAlaGlyArgPro 41
Db      143 CTGGCGGCGAGGAGGCGAGGCGAGCAGCCGCTGCACCGGACGACGGGCTCGAGGCGG 202
Qy      42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56
Db      203 GGACCGCGCGCGCGCGCTGTCTCCCGCGGGTGTCTGCGGGGCGCTGCCCGCGCGCG 262
Qy      57 LeuArgProArgLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeu 76
Db      263 CTGCTGCCG---CTGCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 298
Qy      77 LeuProCysGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
Db      299 CTGCCCGCGGAGCGCGCGCTGCGGTGGCGGCGGCGGTGTCGGGCTCGCGCGCGCG 358
Qy      97 GluAlaLysGluCysAspArgProCysValAlaGlnGlyArgCysAsnProGlyThrGly 116
Db      359 GAGGCCAAGGAATGACCGCGCGGTGTGTCAACGCGCGCGCTGCACACCTTGGCACCGCG 418

```

```

Qy      117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
Db      419 CAGTGCCTCTGCCCCACCGGCTGGTGCGGAGCAATGCCAGCACTCGCGGGCGCGCTTC 478
Qy      136 ----- 136
Db      479 AGGACATCTGTCTCACCCCTATTAATCACAGCTGTTGGAAAGGTGAGGCTGGAGGAACAGT 538
Qy      136 ----- 136
Db      539 TCGAGGCAAGCTTCGGCTACAGAATAAGTTCAAGAGTAACCTGGCGCAACTTGGGCTGTG 598
Qy      136 ----- 136
Db      599 CTCCAAAACCAAAATGAGCGAAAGAGCAAGCTAGAGTCTTTTGGGAAAAATTTTAGCTG 658
Qy      137 -----Arg-LeuThrGlySerSerGlyPheValThrAspGlyProGlyAs 151
Db      659 ACTAAATTTTTCACCGAGAACTAACTGCTCTCTGGATTGTGAACAGATGACCTGGAA 718
Qy      151 nTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetAr 171
Db      719 TTATAAATATAGAAGCAAGTGCATGGCTCATTTGAAGGACAGCCAAATAGAATAATAG 778
Qy      171 gLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGly 191
Db      779 ACTTCGCTTCAACCATTTTGTACAGAAATGATCTGGACCATTTATATATGTTATGATGG 838
Qy      191 YAspSerIleTyrAlaProLeuValAlaPheSerGlyLeuIleValProGluArgAs 211
Db      839 GGAATCAATCTAGCCACCTCTGATTCGCTTCTAGTGGCTCATTTGTTCTCTCAAGAGA 898
Qy      211 pGlyAsnGluThrValProGluValAlaThrSerGlyTyrAlaLeuLeuHisPhePh 231
Db      899 TGGCAATGAGACCGCTCTGAGGTCACTGTCATTCAGGTATGCACTCTGCTGCTGCTGCT 958
Qy      231 eSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysPr 251
Db      959 CAGTGAATGCTGCTTATAATCTGACTGATTTAATAATCACTTCAATTTTACATGTGTC 1018
Qy      251 cAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCly 271
Db      1019 GAATAATTCCTCAGCGGAGGAGGTAAAGACAGTAACAGCAGCAGCGCTGTGAGTG 1078
Qy      271 sGluCysSerGluAsnTrpLysGlyAlaCysAspIlePrCHisCysThrAspAsnCy 291
Db      1079 TGAATGTTCTGAAAACCTGGAAAGGGAGTGGTGTGACATTCCTCACTGTACAGCACTG 1138
Qy      291 sGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSe 311
Db      1139 TGCCTTTCCTCAGCGAGGATCTGTAATCAAGCGATACAGAGGGTGTCTCTGCTTTCC 1198
Qy      311 rAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrAr 331
Db      1199 TCAGTGGAGGGTCTGGATGTTCAATTCCTGCGCAGCTAACCACTCTTTTGGACTCG 1258
Qy      331 gGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAs 351
Db      1259 AGAAGAATATTCGATTTAAAGCTTCCAGAGCTCTCATAAAGCTGTGCTCAATGGAAA 1318
Qy      351 nIleMetTrpValValGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAl 371
Db      1319 TATAATGTGGTGTGTGGCGATATATGTTCAACCATTCAGATTACAGCATGTTTATG 1378
Qy      371 aTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnValValVa 391
Db      1379 GTATGACCTGACTTCTAGGGAATGGCTTCCATTAACCACTCTGTGAACAGTGTGTTGT 1438
Qy      391 lArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyLysI 411
Db      1439 AGATATGTGCTATTCTTTGGCATTTACATAGGATATAATCTACATGTATGGAGGAAAAT 1498

```



QY 411 eAspProThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTr 431  
DB 1499 TGATTCAACAGGGAACGTCACCAATGAGCTGAGAGTATTTCATATTATATGATGATCATG 1558  
QY 431 pValLeuLeuThrProLysAlaLysGluGlyTyrAlaValValGlyHisSerAlaHisIle 451  
DB 1559 GGTATGTGTTAACTCCGAAGCTTAAGATCAGTATGATGAGTGTGGACACTCAGCACACAT 1618  
QY 451 eValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTy 471  
DB 1619 TGTACACTGGCATCTGGCCGCTGGTGTGTCATCTTGGTCACTTCGGTCACTGCCCACTTA 1678  
QY 471 rGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHi 491  
DB 1679 TGGATATATAGCGTGTGCGAGATATGACTTGGAAAAGAACACATGAGGATATATTACA 1738  
QY 491 sThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSerValTyrAspHisArgTh 511  
DB 1739 TACTCAGGGTGTCTTGTGCAAGGGGTATGGCCACAGTAGTGTATTATGATGACAGGAC 1798  
QY 511 rArgAlaLeuTyrValHisGlyGlyTyrIleAlaPheSerAlaAsnLysTyrArgLeuAl 531  
DB 1799 CAAGGCTCTGTACTTCACTGCTGCTACAAAGCTTTCAGCGCCAAACAAATACCGGCTTGC 1858  
QY 531 eAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerAr 551  
DB 1859 AGATGACCTCTACAGATACGATGTGGATCTCAGATGTGGACCATCTTTAAGGACAGCGG 1918  
QY 551 gPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGl 571  
DB 1919 ATTTTTCGGTTACTTGATACAGCTGTGATAGTAGTGGAACCATGCTGGTGTGGAGG 1978  
QY 571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591  
DB 1979 GAACACACAAATGACATTCATGAGCCACCGTCCCAATGCTTCTCTCGGACTTCAT 2038  
QY 591 tAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAs 611  
DB 2039 GGCTTATGACATGTGTGTGACCGATGTGTCAGTGTCTCCAGACCTGAGCTCCATCATGA 2098  
QY 611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGl 631  
DB 2099 TGTCAACAGATTGGCCATTACAGCATCTTGTATACACACAGCCATGTATGTGTTCGGCGG 2158  
QY 631 yPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHi 651  
DB 2159 CTTCAACAGCCCTCTCTCAGTACAGCTTGTGTGCTTTACTCGGAGCAGTGCATGCACA 2218  
QY 651 sArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGl 671  
DB 2219 CCGCAGTGAAGCTGTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACA 2278  
QY 671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSe 691  
DB 2279 GTCGTCTCAGTGTACCTCTCGGAGTGTGCAACTGAAGAACAGCAGAAAGTTAAATC 2338  
QY 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711  
DB 2339 AGAGTGTGTTCTTAAAGAACCTTGCACCATGACAGATGTGACGACACACAGATTGTA 2398  
QY 711 rSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAs 731  
DB 2399 CAGCTGCACAGCCAAATACCAATGACTGCCACTGGTGCAATGATCATCTGTGTCCTGTGAA 2458  
QY 731 nHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAs 751  
DB 2459 CCACAGCTGCACAGAAAGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCAGGATTA 2518  
QY 751 nProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCy 771  
DB 2519 CCCCATGTACTCTGCAATTAAGAAAACCAAGCTGCAGGAGCTGTGCCCTAGACAGAACTG 2578  
QY 771 sGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGl 791

DB 2579 CCACTGGGAGCCCGGAATCAAGAGTGCATCGCCCTCGCCGGAATAATCTGTGGCAATGG 2638  
QY 791 YTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTrpAspAs 811  
DB 2639 CTGGCATTTGGTTGGAACTCGTGTCTGAAATCACTACTCTAAGAGAAATTATGACAA 2698  
QY 811 nAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLy 831  
DB 2699 TGTAAATTTGCTGTAGGAACCAATGCTTTTGGCTTCCTCCATCCACATCCCAAGAA 2758  
QY 831 sValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLe 851  
DB 2759 GGTGGAGTTTGTCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAAGTATGTCCAAGCT 2818  
QY 851 uThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAs 871  
DB 2819 CACTCTGACTCCATGGGTGTGCTCGGAAGATCAATGTGTCTTACTGTGTGGGAGGA 2878  
QY 871 pMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGl 891  
DB 2879 TATGCTCTCCATTCAAAATAGTTTGTGCAAGTGCATGCCATCTGAGCCCAAGTATGCTGG 2938  
QY 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911  
DB 2939 CTTCTGTGGGATCTGTGACAGCCTAGTACTCGGGGATTAAGGCTGCAACCTGCATCAA 2998  
QY 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgTh 931  
DB 2999 CCTCTCAATGGGAGCGTCTGTGAAAGGCTGCAAAACCAAGTGCAGGAGTGCAGGAC 3058  
QY 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951  
DB 3059 ACCATGTGCCCTCGGACAGCGTGTGGGAGTGCACCTAGCAGCAGCTCGAGTGCATGTG 3118  
QY 951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGl 971  
DB 3119 GTGCAGTAACATGAAGCAGTGTGTGGACTCCAAATGCTTACGTGGCCCTCTTCCCTTTGG 3178  
QY 971 yGlnCysMetGluTrpTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCy 991  
DB 3179 CCAGTGTATGAATGGTATAGATGAGCAGCTGCCACCTGAAAATTTGCTCTGGTACTG 3238  
QY 991 sThrCysSerHisCysLeuGluGlnProGlyCysGlyTyrCysThrAspProSerAsnTh 1011  
DB 3239 TACCTGCAGCAATTGCTTGGAGCAGCCAGCGCTGTGGTGTGTACTGATCTCTAGCAATAC 3298  
QY 1011 rGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAl 1031  
DB 3299 TGGAAAAGGAAAATGTATTGAGGCGACGTATAAGGACCTGTGAAGATGCCGTCCAGGC 3358  
QY 1031 aProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSe 1051  
DB 3359 CTCTGCAGGAATGTGTATCCACAGCCCTTCTGAACCTCCAGCATGTGTCTAGAGGACAG 3418  
QY 1051 rArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCy 1071  
DB 3419 CAGATACAACCTGGTCTTTTTCATCTCTCCAGCTTGCAGTGCACGACGACACCAATG 3478  
QY 1071 sIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluTh 1091  
DB 3479 CATCAACAGAGTATCTGTGAGAGTGTGAGGACCTGACCCAGGCGGACGACCTCCGAGAC 3538  
QY 1091 rCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCy 1111  
DB 3539 CTGCATATCTGGCTCTATGTGTGACCGGACTTAATGGAGGCAAAATGTGAGCCATCCAAATG 3598  
QY 1111 sAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGl 1131  
DB 3599 CAATGGGACGCACTACTGTGCAACACCAACCGGCAAGTCTTCTGTACCAACCAAGG 3658  
QY 1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLe 1151



3559	TGTCACGGGGAGCGAGTCCGACGTATGTGAGGCTAGAAAATCGATACCAAGGAACCCCTCT	3719
Qy	uAAGgLyThrCysTyTyThrLeuLeuIleAsePtyrGlnPheThrPheSerLeuSerG1	1171
Db	CAAAGGAACATGCTACTATACCCCTTCTCATTTGACATCAGTTCACTTTAGCCTGTGCCA	3778
Qy	nGluAspAspArgTyTyThrAlaIleAasnPheValAlaThrProAspGluGlnAsnAr	1191
Db	7GAAAGACGACCGCTACTACACAGCCATCAACTTTGTGGCTACTCTCTGATGAACAAACACAG	3838
Qy	gASpLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAl	1211
Db	GGATTGGACATGTTTCATCAATGCTCCAAAACCTTCAACCTCAACATCACTTGGGCCAC	3898
Qy	aSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIl	1231
Db	CAGCTTCCACGCGGAACCCAGACGTGAGAGAGAGTGCCTGTGTTTCAAAAACCAACAT	3958
Qy	eLysGluTyLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIl	1251
Db	CRAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCAACAT	4018
Qy	eThrPhePheValTyValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSe	1271
Db	CACCTTCTTGTGTTTATGTCTAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCCTTCTC	4078
Qy	rGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheIle	1291
Db	CCAGACAGCAACTTCAATGACCTGGTACAGTCTCTCGTACATTTCTTCAGTTGTTTCCT	4138
Qy	uSerLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerAr	1311
Db	CTCGCTCTCTTGTGTGCTGCTGAGTGGTCTGGAAGATCAACAGACAGCTGTTGGCATCCAG	4198
Qy	gArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerVa	1331
Db	CGCGAGAGACCAACTTCTCGGAGATGCAACAGATGGCCAGCGCCCTTTGCTTCTGT	4258
Qy	lAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLysTh	1351
Db	AAACGTGTCCTTGGAAACAGATGAGGAGCCTCTCTGATCTATTGGGGGGAGTATAAGAC	4318
Qy	rValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerVa	1371
Db	TGTTCCAAACCCATTCGACATGGAGCGGTGTTTGTGGCAACAAAGCCGCTGCTCTCTGT	4378
Qy	lPheValArgLeuProArgGlyLeuGlyIleProProGlyGlnSerGlyLeuAl	1391
Db	GTTTGTGAGCTCCCTCGAGCCTGGGTGGCATCCCTCTCTCTGGGAGTCAAGTCTTGC	4438
Qy	aValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyLysGluLysSe	1411
Db	TGTTGGCCAGCGCCTGTGTGACATTTCTCAGCAGATGCCATAGTCTACAAGGAGAAGTC	4498
Qy	rGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle	1429
Db	AGAGGCGGTGAGAAACCGAAGCAGCAGCCCCCTGCACGCTGGGACCTGCGATC	4553
RESULT 7		
AF119821		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
Mus musculus (house mouse)		
Mus musculus		
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
1 (bases 1 to 4313)		
Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,		
Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.		
The mouse mahogany locus encodes a transmembrane form of human		
TITLE		

Db 156 GGACCGCGCGCGGCTGTCTACCGCGGGTGTCTCGCGGGCGCTGCCCGCGCGCGG 215  
Qy 57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 76  
Db 216 CTGCTGCCG---CTGCTCTTTTCGCTGTCTGCTGCTGCCG----- 251  
Qy 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96  
Db 252 CTGCGCGGAGCGCGCGCGCTGCGGTGGCGGCGCGGCTGCGGCTCGGCGCGAGCC 311  
Qy 97 GluAlaLysGluCysAspArgProCysValAsnGlyAlaArgCysAsnProGlyThrGly 116  
Db 312 GAGGCCAAGAAATGTACCGCGCGTGTGTCAACGCGCGCGCTGCAACCCCTGGCAACGCG 371  
Qy 117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyVArgPhe 136  
Db 372 CAGTGGCTGTGCCCGCCAGCGGCTGGGTGGCGGAGCAATGCCAGCATCGGGGGCGGCTTC 431  
Qy 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156  
Db 432 AGACTAACTGGCTCTCTGGATTGTGAAGCTATCCAAATGCAGTGGACCTGGGAATATAAATATAAGACG 491  
Qy 157 LysCysThrTrpLeuIleGluGlnProAsnArgIleMetArgLeuArgPheAsnHis 176  
Db 492 AAGTGCACATGGCTCATTTGAAGGCTATCCAAATGCAGTGGCTAAGGTTAAGATTCAACCAT 551  
Qy 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196  
Db 552 TTTGCTACAGATGTAGCTGGACCATTTATATGTTTATGATGGGAGCTCAATCTACGCA 611  
Qy 197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216  
Db 612 CCTCTGATTGCTGCTTTAGTGGCTCATTTGTCTCGAAGAGATGGCAATGAGACGGCT 671  
Qy 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaIleTyr 236  
Db 672 CTGAGTCACTGCTCACTTCAGGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 731  
Qy 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256  
Db 732 AATCTGACTGGATTAAATATCACTTACAAATTTTGACATGTCTCCGAATAATTTGCTCAGCC 791  
Qy 257 ArgGlyGluCysAlaSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276  
Db 792 CGAGGAGAGTGAAGAGAGTAAACAGCAGCAGCGCTGTTGAGTGTGAATGTTCTGAAGAC 851  
Qy 277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296  
Db 852 TGGAAAGGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 911  
Qy 297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316  
Db 912 GGCATCTGTAATCAAGCATACCAAGAGGTGCTCTGCTTCTCCTCACTGGCAGGGTCTCT 971  
Qy 317 GlyCysSerValProValProAlaAsnGlnSerPheTrpArgGluGluTyrSerAsn 336  
Db 972 GGAATGCTCAATCTCTGCGCAGTACCACTCTTTTGGACTCGAGAGAATATTCGAT 1031  
Qy 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlnAsnIleMetTrpValVal 356  
Db 1032 TTAAGCTTCCCGAGCGCTCTCAATAAGCTGTGGTCAATGGAATATAATATGTGGGTGTT 1091  
Qy 357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376  
Db 1092 GCGCGATATATGTTCAACCATTCAGATTACAGATGGTTTTAGCGTATGACCTGACTTCT 1151  
Qy 377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396  
Db 1152 AGGCAATGGCTTCCACTAAACCATCTCTGTGAACAGTGTGGTGTGAAGATATGGCTATCTCT 1211  
Qy 397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416  
Db 1212 TTGGCATTACATAAGGATATAAATCTACATGTATGGAGGAAAAAATTGAATCAACAGGGAAAC 1271

Qy 417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuLeuThrPro 436  
Db 1272 GTGACCAATGAGCTGAGAGATTTTCATATTCATAATCAATCATGGGTATTTAACTCCG 1331  
Qy 437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456  
Db 1332 AAAGCTAAGGATCAGTATGAGTGGTGGACACTCAGCACACATTTGTACACTGGCATCT 1391  
Qy 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476  
Db 1392 GCGCGTGTGTGTCATGTTGGTCACTCTTCGTCATTTGCCACTATGGATATATAAGGTT 1451  
Qy 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496  
Db 1452 GTGCAGGAATATGACTTTGGAAAGAACACATGGAGTATATACATCTCAGGGTGTCTCT 1511  
Qy 497 ValGlnGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516  
Db 1512 GTGCAAGGGGGTTATGGCCACAGTAGTGTCTTATGATCAGGACCAAGGCTCTGTACGTT 1571  
Qy 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspLeuTyrArg 536  
Db 1572 CATGGTGGCTACAGGCTTTTCAGGCCCAACAATACCGCTTGCAGATGACCTCTACAGA 1631  
Qy 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPheArgTyrLeu 556  
Db 1632 TACGATGTGATATCTCAGATGTGCACCATTTCTTAAGACAGCCGATTTTTCGGTTACTTG 1691  
Qy 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576  
Db 1692 CATACACTGTGATAGTAGTGAGTGNACCTGCTGGTGTGGAGGGAACACACACAATGAC 1751  
Qy 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596  
Db 1752 ACTTCCATGACCCACGGTGCCTTCTCTCCATCAGACTTCATGGCTTATGACATTTGCT 1811  
Qy 597 CysAspArgTrpSerValLeuProArgProAspLeuHisAspValAsnArgPheGly 616  
Db 1812 TGTACCGATGGTCAAGTGTCTCCAGACCTGAGTCCCATCATGATGTCACAGATTTGGC 1871  
Qy 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
Db 1872 CATTCAGCAGTCTTGTACACAGCAGCAGCATGTATGTGTGGCGCTTCAACAGCGCTCCTC 1931  
Qy 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
Db 1932 CTGAGTCACTGCTTGTGCTTTTACCTCGAGCAGTGGCATGCACACCCGAGTGAAGCTGCT 1991  
Qy 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676  
Db 1992 TGTGTGGCAGCAGACCTGATCTCGGTGTCTGTGGACACACAGTCTCTCGATGTACC 2051  
Qy 677 SerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696  
Db 2052 TCCTGGGAGTTGGCAACTGAAGAAACAAGAGTAAATCAGAGTGTCTTTCTTAA 2111  
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
Db 2112 AGAACCTTGTACCATGACAGATGTACAGCAGCAGATTTGTACAGCTGCACAGCCAAT 2171  
Qy 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
Db 2172 ACCAATGACTGCCACTGGTGCATGATCACTGTCTCTGTGAACACACAGCTGCACAGAA 2231  
Qy 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
Db 2232 GCGCAGATCTCCATTGCCAATGAGATGTGCGCCCAAGGATAACCCCATGTACTACTGCG 2291  
Qy 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776  
Db 2292 AATAGAAACACAGCTGCAGGAGCTGTGCCCTAGACCAAGACTGCCAGTGGAGCCCGG 2351

QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisIleuValGly 796  
 DB 2352 AATCAAGAGTGCATCGCCCTGCCGGAATAATATCTGTGGAATGGCTGGCAATTTGGTTGA 2411  
 QY 797 AsnSerCysLeuIlysIleThrAlaIysGluAsnTyrAspAsnAlaIysLeuPheCys 816  
 DB 2412 AACTCGTGTGTAATCACTACTGTAGGAGAAATATGACAAATGCTAAATGGCTGT 2471  
 QY 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrGlnIysIysValGluPheValLeu 836  
 DB 2472 AGGAACCAATGCCCTTTTGGCTTCCCTCACATCCAGAGAAGGTGGAGCTTGTCTT 2531  
 QY 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856  
 DB 2532 AACGACCTTCGTAATGCAATCATCTCAAGATATGTCAGCTCACTCTGACTCCATGG 2591  
 QY 857 ValGlyLeuArgIysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876  
 DB 2592 GTTGGTCTTCGGAAGATCAATGTGCTTACTGTGTGCTGGAGGATATGCTCCATTCACA 2651  
 QY 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
 DB 2652 AATAGTTTGTGTCAGTGGATGTCATCTGAGCCCGAGTGTGCTGCTTCTGTGGATCTTG 2711  
 QY 897 SerGluProSerThrArgGlyLeuIysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
 DB 2712 TCAGAGCTTAGTACTCGGGGATTAAGGCTGCAACCTGTCATCAACCTCTCAATGGCAGC 2771  
 QY 917 ValCysGluArgProAlaAsnHisSerAlaIysGlnCysArgThrProCysAlaLeuArg 936  
 DB 2772 GTCTGTGAAGGCTGCAACCAACACAGTGCACAGCAGTGCAGGACCACTGTGCCCTGGG 2831  
 QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956  
 DB 2832 ACAGCTGTGGCGAGTGCACACTAGCAGCAGCTCGAGTGTATGTGTGTCAGTAACTGAAG 2891  
 QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyClnCysMetGluTrp 976  
 DB 2892 CAGTGTGTGACTCCCAATGCTACGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGG 2951  
 QY 977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
 DB 2952 TATACATGAGCAGCTGCCACCTGAAATGTCTGTGGTACTGTACCTGCGACCATTCG 3011  
 QY 997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyIysGlyCys 1016  
 DB 3012 TTGGACGACGAGCTGTGTTGTTGCTACTGATCTAGCAATACTGGGAAGAAATGT 3071  
 QY 1017 IleGluGlySerTyrIysGlyProValIysMetProSerGlnAlaProThrGlyAsnPhe 1036  
 DB 3072 ATTGAGGCGAGCTATAAAGACCTGTGAAGATGCCGTACAGGCCCTCTGCGAGAAATGT 3131  
 QY 1037 TyrProGlnProLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056  
 DB 3132 TATCCACAGCCCTTCTGAACTCCAGATGTCTTAGAGGACAGCAGATACAACTGGTCT 3191  
 QY 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
 DB 3192 TTCAATTCAGTCTCCAGTTCGCAACGAGACAGCAATGCAATGCATCAACACAGAGTGT 3251  
 QY 1077 CysGluIysCysGluAsnLeuThrThrGlyIysHisCysGluThrCysIleSerGlyPhe 1096  
 DB 3252 TGTGAGAATGTGAGAGCTGACCCAGGCAAGCACTGGAGACCTGCAATATCTGGCTTC 3311  
 QY 1097 TyrGlyAspProThrAsnGlyIysCysGlnProCysLysCysAsnGlyHisAlaSer 1116  
 DB 3312 TATGGTGAACCCGACTAATGAGGCAAAATGTAGCCATGCAAGTGCATGGGCAAGCACTA 3371  
 QY 1117 LeuCysAsnThrAsnThrGlyIysCysPheCysThrThrLysGlyValIysGlyAspGlu 1136  
 DB 3372 CTGTGCAACACCAACCCGCAAGTCTTCTGTACCAACCAAGGTGTCAAGGGGACGAG 3431  
 QY 1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156

DB 3432 TGCAGACTATGTAAAGTAGAAAAATCGATACCAAGGGAACCTCTCAAGAAACATGCTAC 3491  
 QY 1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176  
 DB 3492 TATACCCCTTCTCATTTGACTATCAGTTTACCTTTAGCCTGTCCCGGGAGACGACCTAC 3551  
 QY 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196  
 DB 3552 TACACGCCATCAACTTTTGTGGCTACTCTCTGATGACAAACAGGGATTTTGCACTGTT 3611  
 QY 1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaIleAsnPheSerAlaGly 1216  
 DB 3612 ATCAATGCCCTCCAAAAATTCACCTCAACATCACCTGGGCCACAGCTTCCAGCGGA 3671  
 QY 1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGlyIysLysAsp 1236  
 DB 3672 ACCCAGACTCGAGAAGAGTGCCTGTGTTTCAAAACCAACATCAAGGAATCAAGAT 3731  
 QY 1237 SerPheSerAsnGluIysPheAspPheArgAsnHisProAsnIleThrPheValTyr 1256  
 DB 3732 AGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACTTCTTTGTTAT 3791  
 QY 1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276  
 DB 3792 GTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCCCTTCTCCAGCACACCAACTT 3851  
 QY 1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuVal 1296  
 DB 3852 ATGACCTGTGTACATTTCTCGTACTTCTTCAGTTGTTTCTCTCGCTGCTTCTGGTG 3911  
 QY 1297 AlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeu 1316  
 DB 3912 GCTGCACTGTCTGGAAGATCAAGCAGAGCTGTGGGCATCCAGCGGAGAGACACTT 3971  
 QY 1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336  
 DB 3972 CTTGGGAGATGCAACAGATGGCCGCGCCCTTCTGTCTGTAAACGTTGCTTGGAA 4031  
 QY 1337 ThrAspGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356  
 DB 4032 ACAGATGAGAACTCTCTGATCTCATTGGGGAGATATAAGACCGTTCTTCAAGCCATT 4091  
 QY 1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376  
 DB 4092 GCCCTGAGCCCTGCTTGGTAAACAAAGCCGAGTCTCTCTCTATTCTGTAGGCTCCCT 4151  
 QY 1377 ArgGlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396  
 DB 4152 CGAGGACTGGGAGAAATCCCTCTCTGTTGTCAGTCAAGTCTCGCTGTGGCCAGTCCCTG 4211  
 QY 1397 ValAspIleSerGlnGlnMetProIleValTyrLysGlyLysSerGlyAlaValArgAsn 1416  
 DB 4212 GTGACATTTCTCAGCAGATGCCAAATAGTGTACAAGGAGAGTCAGGAGCTGTAAAGAAC 4271  
 QY 1417 ArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429  
 DB 4272 CGAAGCAGCAGCGCCCTGCACAGCTGGAACCTGCATT 4310

## RESULT 8

AF106861 3819 bp mRNA linear PRI 06-JAN-1999  
 LOCUS Homo sapiens attractin-2 (ATRN) mRNA, complete cds.  
 DEFINITION Homo sapiens attractin-2 (ATRN) mRNA, complete cds.  
 ACCSSION AF106861  
 VERSION AF106861.1 GI:4093195  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 3819)  
 REFERENCE Duke-Cohan,J.S., Gu,J., Freeman,G.J. and Schlossman,S.F.  
 AUTHORS Cloning of cDNA for attractin-2, identical with that of attractin  
 TITLE



1321 DB CAGTATGCGAGTGGTGGGCACTCTGCACACATGTTTACACTGAAGAATGGCGAGTGGTC 1380  
461 QY MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480  
1381 DB ATGCTGGTCACTCTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGATAT 1440  
481 QY AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500  
1441 DB GATTTGGATAAAGAACACATGGAGATATATTACACACCCAGCGGTGCCCTTGTGCAAGGGGT 1500  
501 QY TyrGlyHisSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
1501 DB TACGGCCATAGACGTGTTTACGCCATAGACACCGAGCGCCCTATATAGTTTATGGTGGCTAC 1560  
521 QY LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540  
1561 DB AAGGCTTTCAAGTGCCTAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGGAT 1620  
541 QY ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrIleuHisThrAlaVal 560  
1621 DB ACCCATGTGGACCATCTTAAAGGACAGCCGATTTTCCGTTACTTGCACACAGCTGTG 1680  
561 QY IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
1681 DB ATAGTCAGTGGACCATGCTGGTGTGTTGGGGGAAACACACACATGACACATCTATGAGC 1740  
581 QY HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600  
1741 DB CATGGCGCCAAATGCTTCTCTTCAAGATTTTCAATGGCTTATGACATTCCTGTGACCGCTGG 1800  
601 QY SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620  
1801 DB TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCACAGTC 1860  
621 QY LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640  
1861 DB TTACAAACAGCACCATGATGTGTGCGGTGGTTTCAATAGTCTCTCTCTCCTCAGCAGCATC 1920  
641 QY LeuValPheThrSerGlnGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660  
1921 DB CTGGTATTCACTCGGAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTATTAGCAGCA 1980  
661 QY GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680  
1981 DB GGACCTGGTATTGCGTGTGTGGAAACACAGGCTCGTCTCAGTGTATCTCGTGGGCGCTG 2040  
681 QY AlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700  
2041 DB GCAACTGTATGAACAAGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAGAACTCTTGAC 2100  
701 QY HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720  
2101 DB CATGACAGATGTGACAGCACACAGATTTTACAGCTGTACAGCAACACCAATCACTGTC 2160  
721 QY HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
2161 DB CACTGTGCAANTGACCATTTGTGTCACAGAACCCACAGCTGTCTCAGAGGGCCAGATCTCC 2220  
741 QY IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760  
2221 DB ATTTTGTAGTATGAGATTTGCCCAAGGATATACCCCATGTACTGTATTAACAGAGACC 2280  
761 QY SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780  
2281 DB AGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAGTGGAGGCCCGGAATCAGGAGTGC 2340  
781 QY IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800  
2341 DB ATTGCCCTGCCCGAAATATCTGTGGCATGCTGGCATTTGGTTGGAAACTCATGTTTG 2400  
801 QY LysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820

2401 DB AAAATTACTACTGCCAAGGACAAATTATGACAATGCTAAATGTTTCTGTAGACACCAAT 2460  
821 QY AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840  
2461 DB GCCCTTTGGCTTCTCTTACCAACCCAGAGAGTAGAATTTGCTCTTAAGCAGCTGCCA 2520  
841 QY IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860  
2521 DB ATAATGAGTCACTCTCAGACGATGTCAGACTCACCTTAACCCATGGGTGGCCCTTCGG 2580  
861 QY LysIleAsnValSerTyrTrpCysTrpGluAspMetSerTrpPheThrAsnSerLeuLeu 880  
2581 DB AAGATCAATGTGCTCTACTGTCGCTGGGAAGATATGTCCCCCAITTTACAAATAGTTACTA 2640  
881 QY GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
2641 DB CAGTGGATGCGCTCTGAGCCGAGTGTGCGAATCTGTGGAATTTTATCAGAACCCAGT 2700  
901 QY ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
2701 DB ACTCGGGGACTGAAGCTGCAACTGCATCAACCCACTCAATGCTAGTGTCTGTGAAGG 2760  
921 QY ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
2761 DB CCTGCAAAACACAGTGTAAAGCAGTGCAGCAGACCATGTGCTTGAAGCAGCATGTGA 2820  
941 QY AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960  
2821 DB GATTGACACAGCGAGCTCTGAGTGCATGTGTGCGACCAATGAGCAGTGTGTGGAC 2880  
961 QY SerAsnAlaTyrValAlaSerPhePheGlyGlnCysMetGluTrpTyrThrMetSer 980  
2881 DB TCCAAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACATGAGC 2940  
981 QY ThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnPro 1000  
2941 DB ACCTGCCCCCTGAAATTTGTCAGGCTACTGACCTGTAGTCTGTAGTCTTGTGGAGCAACA 3000  
1001 QY GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020  
3001 DB GGCTGTGGCTGGTGTACTGATCCAGCAATFACTGCGCAAGGAAATGCTAGAGGGTTC 3060  
1021 QY TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040  
3061 DB TATAAAGACACAGTGAAGTGCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCC 3120  
1041 QY LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060  
3121 DB CTGCTCAATTCAGCATGTCTTAGAGGACAGACAGATACACTGGTCTTTCATTCACTGT 3180  
1061 QY ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080  
3181 DB CCAGCTTGCCAATGCAACCGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGTGT 3240  
1081 QY GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100  
3241 DB GAGAACCTGACCAAGCAAGCAACTGCGAGACTGTCATATCTGGCTTCTACGGTGTATCCC 3300  
1101 QY ThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120  
3301 DB ACCAATGGAGGAAATGTCTAGCCATGCAATGGGCAACGCGTCTCTGTGCAACACC 3360  
1121 QY AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140  
3361 DB AACACGGGCAAGTGTCTTGCACACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGT 3420  
1141 QY GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160  
3421 DB GAGGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGGAAACATGTTATTATATCTCTCT 3480  
1161 QY IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180  
3481 DB ATTGACTATCAGTTACCTTTAGTCTATCCCCAGGAAGATGATCGCTATTACAGACTATC 3540





Db 1366 GGCCGAGTGGTCATGCTGCTCATCTTTGGTCACTGCGCTCTCTATGGATATATAGCAAT 1425  
QY 477 valGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496  
Db 1426 GTGCGAGAAATGATGATTGGATAGAAACACATGGAGTATATTACACACCCAGGGTGCCTT 1485  
QY 497 valGlnGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516  
Db 1486 GTGCAGGGGGTTACGGCCATAGCAGTGTTCAGCCATAGAGACAGGCGCCCTATACGTT 1545  
QY 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspLeuTyrArg 536  
Db 1546 CATGGTGGCTTACAAAGGCTTTCAGTCCAAATAGTACCGCTTGCAGATGATCTCTACCGA 1605  
QY 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPheArgTyrLeu 556  
Db 1606 TATGATGTGATACCCAGATGTGGACCTTTTAAAGACAGCGGATTTTTCGTTACTTG 1665  
QY 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576  
Db 1666 CACACAGCTGTGATAGTAGTGGAACCATGCTGGTGTGTTGGGGGAAACACACACAATGAC 1725  
QY 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596  
Db 1726 ACATCTATGAGCCATGGCCCAATGCTCTCTTCAGATTTTCATGGCTTATGACATGGCC 1785  
QY 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616  
Db 1786 TGTGACCGCTGTGTCAGTGTCTCCACACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845  
QY 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
Db 1846 CATTCAGCAGCTCTTACACACAGCACCATGATGTGTTCCGTTGGTTTCAATAGTCTCCTC 1905  
QY 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
Db 1906 CTGAGCGCATCTCGTATTTCACCTCGGAACAGTGTGATCGCAGTCCGAGTGAACCCGCT 1965  
QY 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676  
Db 1966 TGTTTAGCAGCAGGACCTGGTATTCGTTGTTGTTGGAACACAGGCTCGTCTCAGTGTATC 2025  
QY 677 SerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696  
Db 2026 TCGTGGCGCTGCAACTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCCAA 2085  
QY 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
Db 2086 AGAACTCTTGACCATGACAGATGTGACACAGCACAGATTTGTACAGCTGCACGCCAAC 2145  
QY 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
Db 2146 ACCAATGACTGCCACTGGTGCAATGCCATTCCTGTCCTCCAGGAACACAGCTGCTCAGAA 2205  
QY 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
Db 2206 GGCCAGATCTCCATTTTATGATATGAGATTTGCCCCAGGATAACCCATGTAATCTGT 2265  
QY 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776  
Db 2266 AACAGAAGACACAGCTGCAGGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGGCCCGG 2325  
QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796  
Db 2326 AATCAGAGTGCAATGCCCTGCCCGGAAATATCTGTGGCATTTGGTGGGA 2385  
QY 797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
Db 2386 AACTCATGTTTGAATTAATACTGCAAGGAGAAATATGACAATGCTAAATTTGTTCTGT 2445  
QY 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrGlnLysIleValGluPheValLeu 836  
Db 2446 AGGAACCAATGCCCCCTTTTGGCTTCTCTTACAAACCCAGAGAAGGTAGATTTTGTCTT 2505

QY 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856  
Db 2506 AAGCAGCTGCGAATATATGCACTCATCTCAGAGCATGTCNAAGCTCACCTTAACCCATGG 2565  
QY 857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerPropheThr 876  
Db 2566 GTCGGCTTCGGAAGATCAAAGTGTCCTCTGCTGCTGGGAAGATATGTCCTCCATTTACA 2625  
QY 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
Db 2626 AATAGTTTACTACAGTGATGCGCTCGAGCCAGTGATCTGGATCTGTGGAATTTTA 2685  
QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
Db 2686 TCAGAACCCAGTACTCGGGGACTGAAGCTGCAACCTGCACTCAACCTCAATGGTAGT 2745  
QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
Db 2746 GTCTGTGAAAGGCTGCAAAACCAACAGTCTAAGCAGTCCGCGACACCATGTGCTTGAGG 2805  
QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956  
Db 2806 ACAGCATGTGGAGATTGCCACAGCGGAGCTCTGAGTGCATGTGTGCGACACATGAG 2865  
QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrp 976  
Db 2866 CAGTGTGTGACTCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAAATGG 2925  
QY 977 TyrThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
Db 2926 TATACGATGAGCACTGCCCCCTGAAATTTGTTCAGGCTACTGTACTGTACTGTCTG 2985  
QY 997 LeuGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCys 1016  
Db 2986 TTGGAGCAACCGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGAAATGTC 3045  
QY 1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036  
Db 3046 ATAGAGGGTTCTTATAAAGGACCAAGATGCTTCCGAAAGCCCTCAGGAAATTTTC 3105  
QY 1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056  
Db 3106 TATCCACAGCCCTGTCTCAATTCAGCATGTGTCTAGAGGACAGACATACACTGGTCT 3165  
QY 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
Db 3166 TTCATTCACTGTCCAGCTTCCCAATGCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225  
QY 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096  
Db 3226 TGTGAGAAGTGTGGAACCTGACCAACGGAACAGCATGCGAGACCTGCATATCTGGCTTC 3285  
QY 1097 TyrGlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116  
Db 3286 TACGGTGATCCCAACCAATGGAGGGAATGTCCAGCCATGCAAGTGCATGCGGCGCTCT 3345  
QY 1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136  
Db 3346 CTGTGCAACACCAACACGGGCAAGTGTCTTGTGCACCAACGAGGCGGTCAAGGGGACGAG 3405  
QY 1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156  
Db 3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCTCTCAGAGAAACATGTTAT 3465  
QY 1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176  
Db 3466 TATACTCTTCTTATGACTATCAGTTACCTTTAGTCTATCCCGAAGAGATGCTGCTAT 3525  
QY 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196  
Db 3526 TACACACTATCAATTTTGTGGCTACTCTCTGCGAACAAACAGGGATTTGGACATGTTTC 3585



QY 1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrThrAlaAlaSerPheSerAlaGly 1216  
 Db 3586 ATCAATGCTCCCAAGAAATTCACCTCAACATCACTGGCTGCCAGTTCTCAGCTGGA 3645  
 QY 1217 ThrGlnAlaGlyGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236  
 Db 3646 ACCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAAGAGTACAAAGAT 3705  
 QY 1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256  
 Db 3706 AGTTCTCTTAATGAGAAGTTTCATTTGCAACACCACCAAAATATCACTTCTCTTTAT 3765  
 QY 1257 ValSerAsnPheThrTyrProIleLysIleGlnIleAlaPheSerGln 1272  
 Db 3766 GTCAAGTAATTCACCTGGCCCAATCAAAATTCAGGTGCAAACTGAACAA 3813

RESULT 10  
 AF034957  
 LOCUS  
 DEFINITION Homo sapiens secreted T cell activation protein Attractin  
 ACCESSION AF034957  
 VERSION AF034957.1 GI:3676346  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 3597)  
 Duke-Cohan,J.S., Gu,J., McLaughlin,D.F., Xu,Y., Freeman,G.J. and  
 Schlossman,S.F.  
 Attractin (DPPT-L), a member of the CUB family of cell adhesion and  
 guidance proteins, is secreted by activated human T lymphocytes and  
 modulates immune cell interactions  
 Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)  
 98409658  
 9736737  
 2 (bases 1 to 3597)  
 Duke-Cohan,J.S., Gu,J., Ao,Z., McLaughlin,D.F., Freeman,G.J. and  
 Schlossman,S.F.  
 Direct Submission  
 Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS,  
 Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,  
 USA

FEATURES  
 Source  
 Location/Qualifiers  
 1..3597  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 1..3597  
 /gene="attractin"  
 1..3597  
 /gene="attractin"  
 /note="secreted T cell activation protein; Method:  
 conceptual translation with partial peptide sequencing;  
 DPPT-L"  
 /codon\_start=1  
 /product="Attractin"  
 /protein\_id="AAC61902.1"  
 /db\_xref="GI:3676347"  
 /translation="WVAAATEARLRRTAATAALAGSGGPHCVNGRCNPGTQC  
 VCPAGVQCCQHCGRFRLTGSGFVIDPGNYKYTKCTWLEGGPNRLMRFPNH  
 FATECSWDLHYDGDSTIYAPLVAAPSGLIPEERDNETVPEVATSGYALHFFSDA  
 AYNLTFNITGYDFMCPNNCSGRGCKISNSDTVECECSNWKGEACDI PHCTDNCG  
 FPHRGICNSDVRGSCFSDMQPGCVFVPANQSFRTREYSNLKLPRAHKA VVNG  
 NIMVVGVYNSHDYNNMVLAYDLASRWLPLNRVNNVVYRYSGLALYKDKIYMYG  
 GKIDSGYNTRELRFPHLHNSWLLTPKAEQYAVGVGSAHIVTLKNGRVVMLVIFG  
 HCLPYGISNVEYDLDKNTWSILHTQALVQGGVGHSSVTHRRFRLYLVHGKYAFS  
 AKHRYLADLDYDVTQMTILKDSFRFRYLHTAVIVSGTLMVYFGVTHNDTSMHSG  
 AKCFSDFMAYDIACDRSVLPRDSTMTSDTALPAVLHNSITMYVFGFNLSLSDI  
 LVFTSGQCDNAHSEACLAAGPGICRVNWTGSSQICISWALATDEQEKLECFPSKRT  
 LDHRCDOHDCYCTANTNDCHWCNHCVPNHSCEGQISIFRYENCKDNPMWYC  
 NKKTSRSCSALLDQNCQWEPNRQECIALPENICIGIWHLVGNSCLKITITAKENYDNAKL

FCRNHALASLTQKKVEFLKQLRIMOSSQMSKLTLPWVGLRKINVSVCWEDM  
 SPTNLSLQMPSEPDSADGFCGLSEPTSLGLKAATCINPLNGSVCEPANHSAKQCR  
 TPCALARTACDSTGSGSECMWCSNMKVAVASVFPFGQCMETWTASTCPFPNCS  
 QCTCSHCLBQPCGCTPSNTGKGKCIEGSYKGVKMPSPQATGNFYFQPLNNSM  
 CLDSRVNYSFIHCPACQNGSHSKINOSTCEKCNLTGKHCTCETCISGYPGLTNGG  
 KCQPCXONGHASLNTNKGKCTTKGVKDECOLCEVENRYOGNPLRGTCTCYTLLID  
 YQTFISLSQDDDRYTAINFVATPDEQNRDLDFINASKNFNLNITWASFSAGTQAG  
 EEMPVSVSKTNIKKEYKDSFNEKDFRHNHNPITPFVIVSNFTWPIKIQVQTEQ"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 3597  
 Score: 6612.00 Matches: 1191  
 Percent Similarity: 93.72% Conservatives: 2  
 Best Local Similarity: 93.56% Mismatches: 5  
 Query Match: 83.54% Indels: 76  
 DB: Gaps: 1

US-09-787-097-12 (1-1429) x AF034957 (1-3597)

QY 1 MetValAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAla 20  
 Db 1 ATG3TGGCCGACGCGCGCACTGAGGCAAGGCTGAGGAGGAGCGCGGCGACGGCA 60  
 QY 21 AlaLeuAlaGlyArgSerGlyProHisTyrAspTyrAspValThrArgAlaGlyArg 40  
 Db 61 GCGCTCGCGGACGAGCGCGCGCGCGCAC----- 90  
 QY 41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60  
 Db 90 ----- 90  
 QY 61 LeuLeuLeuLeuLeuLeuProProLeuLeuLeuLeuLeuLeuProCysGlu 80  
 Db 90 ----- 90  
 QY 81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100  
 Db 90 ----- 90  
 QY 101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120  
 Db 91 -----TGTTCAACGGCGGTGCGTGCACACCTCGCACCGCGCGCTGCTGC 138  
 QY 121 ProAlaGlyTyrValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140  
 Db 139 CCGCGCGCTGGTGGCGGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAATCGGA 198  
 QY 141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTyr 160  
 Db 199 TCTTCTGGTTTGTGACAGATGGACCTGGAAATTTATAAATACAAAACGAAGTGCACGTGG 258  
 QY 161 LeuIleGluGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180  
 Db 259 CTCATTGAAGACACGCCAAATAGAAATAGAGACTTCGTTTCAATCACTTTTGTACAGAG 318  
 QY 181 CysSerTyrAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200  
 Db 319 TGTAAGTTGGACCACTTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGTTGCT 378  
 QY 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGlnThrValProGluValVal 220  
 Db 379 GCATTTAGTGGCTCATGTGTTCTCGAGAGAGATGGCAATGAGACTGTCCCTGAGGTGTT 438  
 QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240  
 Db 439 GCCACATCAGTTATGCTTGTCTGCACTTTTATGATGCTGCTTATATTTGACTGGA 498  
 QY 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260  
 Db 499 TTTAATATTACTTACAGTTTGTATGTGTCCAAATAACTGCTCAGCGCAGAGAGATGT 558  
 QY 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTyrLysGlyGlu 280

Db 559 AGATCAGTAATAGCAGCGTACTGTTGAATGTAATGTTCTGAAACATCGGAAGGTGAA 618  
QY 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300  
Db 619 GCATGTGACATTCCTCACTGATCAGACAACTGTGGTTTCTCCTCATCGAGGCATCTGCAAT 678  
QY 301 SerSerAspValArgGlyCysSerCysPheSerAspTyrGlnGlyProGlyCysSerVal 320  
Db 679 TCAGTCATGTCAGAGATGCTCCTGCTTCTCAGACTGCAGGGTCTCTGGATGTTCAAGTT 738  
QY 321 ProValProAlaAsnGlnSerPheTyrThrArgGluGluTyrSerAsnLeuLysLeuPro 340  
Db 739 CCTGTACACAGTAACCAAGTCATTTGGACTCGAGAGGAATATTTAACTTAAGCTCCCC 798  
QY 341 ArgAlaSerHisIleValValAlaGlnGlyAsnIleMetTyrValValGlyGlyTyrMet 360  
Db 799 AGAGCATCTCAATAAGCTGTGTCAATGGAACAATATATGTGGTGTGGAGGATATATG 858  
QY 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTyrLeu 380  
Db 859 TTCAACCACTCAGATTATAACATGGTTCATAGCGTATGACCTTGCTTAGGGAGTGGCTT 918  
QY 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400  
Db 919 CCACATAACCGTCTGTGAACAATGGTGTGTAGATATGGTCATTTCTTTGGCATTTATAC 978  
QY 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420  
Db 979 AAGGATATAATTTACATGTATGGAGGAAAAATGATTCACTGGGAAATGTGACCAATGAG 1038  
QY 421 LeuArgValPheHisIleHisAsnGluSerTyrValLeuLeuThrProLysAlaLysGlu 440  
Db 1039 TTGAGAGTGTTCACATTCATATAGTCACTGGTGTGTGACCCCTAAGGCAAGGAG 1098  
QY 441 GlnTyrAlaValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460  
Db 1099 CAGTATGCAAGTGTGGGCACTCTGCACACATTTTACACTGAAGAAATGGCCGAGTGTCT 1158  
QY 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480  
Db 1159 ATGCTGGTCATCTTTGGTCACCTGCCCTCTATGGATATATAAGCAATGTGCAGGAATAT 1218  
QY 481 AspLeuAspLysAsnThrTyrSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500  
Db 1219 GATTTGGATAAGAACACATGGAGTATATTACACACCCAGGGTGCCCTGTGCAAGGGGT 1278  
QY 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
Db 1279 TAGGGCCATAGCAGTGTTTACGACCATAGCAGGGCCCTATACGTTTCATGGTGGCTAC 1338  
QY 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540  
Db 1339 AAGGCTTTTCAGTGCCCAATAGTACCGGCTTGAGATGATCTTACCGATATGATGGAT 1398  
QY 541 ThrGlnMetTyrThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560  
Db 1399 ACCAGATGTGGACCAATCTTAAGGACAGCCGATTTTCCGTACTTCGACACAGCTGTG 1458  
QY 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
Db 1459 ATAGTGAGTGAACCACTGTGGTGTGGGGGAAACACACACAATGACACATCTATGAGC 1518  
QY 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTyr 600  
Db 1519 CATGGCCCAATGCTTCTTCAGATTTTCATGGCCCTATGACATTTGCCCTGTGACCGCTGG 1578  
QY 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620  
Db 1579 TCAGTGCTTCCAGACCTGATTC - CACCATGATGTCAACAGATTTGGCCATTCACGACAGT 1637  
QY 620 IleuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIle 640

Db 1638 CTTACACACAGCACCACCATGTATGTGTTCCGTGTTTCAATAGTCTCTCTCAGCGACAT 1697  
QY 640 eLeuValPheThrSerGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAl 660  
Db 1698 CCTGGTATTACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTGTTTAGCAGC 1757  
QY 660 aglyProGlyIleArgCysValTyrAsnThrGlySerSerGlnCysIleSerTyrAlaLe 680  
Db 1758 AGGACCTGGTATTCTGGTGTGTGGAACAACAGGGTGTCTCAGTGTATCTCGTGGCGCT 1817  
QY 680 uAlaThrAspGluGlnGluLysIleLysSerGluCysPheSerLysArgThrLeuAs 700  
Db 1818 GGCACATGATGAACAAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGA 1877  
QY 700 pHsAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCy 720  
Db 1878 CCATGACAGATGTGACCAGCACAGATTGTACAGCTGCACAGCCACCAACCAATGACTG 1937  
QY 720 sHisTyrCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSe 740  
Db 1938 CCACTGGTGAATGACCATTTGTGTCCCGAGAAACCAAGCTGTCTCAGAAAGCCAGATCTC 1997  
QY 740 rIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysTh 760  
Db 1998 CATTTTAGTATGAGAAATTTGCCCAAGGATACCTTATGACTACTGTATACAAAGAGAC 2057  
QY 760 rSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTyrProArgAsnGlnGluCy 780  
Db 2058 CAGCTGCAGGAGTGTGCCCTGGACCAAGAACTGCCAGTGGGAGCCCGGAATCAGAGTG 2117  
QY 780 sIleAlaLeuProGluAsnIleCysGlyIleGlyTyrHisLeuValGlyAsnSerCysLe 800  
Db 2118 CATTCCTCCCGAAATATCTGGCAATGGCTGGCAATTTGGTTGGAAACTCATGTTT 2177  
QY 800 uLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAs 820  
Db 2178 GAAATTTACTTCTGCCAAGGAGAAATATGACATGCTAAATGTTCTGTAGGAACACCAA 2237  
QY 820 nAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuAr 840  
Db 2238 TGCCCTTTTGGCTTCTCTTACAAACCCAGAGAGGAGTAGAATTTGCTTAAAGCAGCTGG 2297  
QY 840 gIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTyrValGlyLeuAr 860  
Db 2298 AATPATGCAGTCTATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTGCGCTTCG 2357  
QY 860 gLysIleAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsnSerLeuLe 880  
Db 2358 GAAGATCAATGTGTCTTACTGGTGTGGGAAGATATGTCCCAATTTACAAATAGTTTACT 2417  
QY 880 uGlnTyrMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSe 900  
Db 2418 ACAGTGATGCCCTCTGAGCCAGTGTGATGCTGTGATTTCTGTGGAATTTTATCAGAACCCAG 2477  
QY 900 rThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluAr 920  
Db 2478 TACTCGGGGACTGAAGGCTCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAG 2537  
QY 920 gProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGl 940  
Db 2538 GCCTGCAAAACACAGTGTGAAGCAGTGCAGGACACCAATGTGCTTGGAGACAGCATGTGG 2597  
QY 940 yAspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAs 960  
Db 2598 AGATTGACACAGCGCAGCTCTGAGTGCATGTGTGTGCAGCAACATGAAGCAGTGTGTGGA 2657  
QY 960 pSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSe 980  
Db 2658 CTCCAATGCCATGTGCGCTCTCTCCCTTTTGCCCAAGTGTATGGAATGTTATACAGTAG 2717  
QY 980 rThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPr 1000  
Db 2718 CACCTGCCCCCTGAAAAATTTGTTTCAGGCTACTGTACTGTAGTCAATTCGTTGGAGCAACC 2777





Db 2401 CAAGAGTGCATCGCCTGCCGAAACATCTGTGTGGCTGGCAATTTGGTGGAAAC 2460  
Qy 798 SerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArg 817  
Db 2461 TCGTGTCTGAAATCACTACTCTAAGGAAATATGCAATGCTAAATTTGCTCTGTAGA 2520  
Qy 818 AsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLys 837  
Db 2521 ACCACAATGCCTTTTGGCTGCTAGTCCAGAGAGAGTGGAGTTGTCTTTAAG 2580  
Qy 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLysLeuThrLeuThrProTrpVal 857  
Db 2581 CAGCTTCGATTATGCAATCATCTCAGAGCAGCTCCAAAGCTCACTCTGACTCCATGGGTT 2640  
Qy 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTyrGluAspMetSerProPheThrAsn 877  
Db 2641 GGTCTTCGGAAGATCAATGATCTTACTGGTCTGGAGGATATGCTCCATTCACAAAT 2700  
Qy 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
Db 2701 AGTTTGTCTCAGTGGATGCATCTCAGCCAGCTGATGCCGGCTTCTGTGGATCTGTCA 2760  
Qy 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917  
Db 2761 GAGCCTAGTACTGGGGACTGAAGGCTGCACTGCATCAACCCACTCAATGGAAGTGT 2820  
Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937  
Db 2821 TGTGAAGSCCTGCAACACAGTCCAAAGCAGTCCGACACCTGTGCTGCCGCGACA 2880  
Qy 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957  
Db 2881 GCATGTGTGAGTGCACCAAGCAGCAGCTCCGAAATGCATGTGGTGAGCAACATGAAGCAG 2940  
Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977  
Db 2941 TCGCTAGACTCAATGCCCTAGTACCTCTTCTTCCTTCGGCCAGTGTATGGAGTGGTAT 3000  
Qy 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
Db 3001 ACAATGAGCAGCTGCCACCTGAAATATGCTCTGCTACTGTACCTGAGCCATGTGCTG 3060  
Qy 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyCysGlyLysCysIle 1017  
Db 3061 GAGCAGCCAGCTGTGGTGGTGTACTGATCTAGCAATCTGGAAGGCGAAATGTATT 3120  
Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
Db 3121 GAGGGAAGCTATAAGGACCCGCTGAAGATGCCCTTCACATGCTCTACAGGAATGTGTAC 3180  
Qy 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057  
Db 3181 CCACAGCCCTTCTGAACCTCCAGCATGTGTCTAGAGGACAGACATCAACTGGTCTTTT 3240  
Qy 1058 IleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCys 1077  
Db 3241 ATTCATGTCCAGCTTCCAGTGTCAATGATGACACAGCAAGTGCATCAACGAGCACTGT 3300  
Qy 1078 GluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyr 1097  
Db 3301 GAGAAGTGTGAGGACCTGACCAAGCAGCAGCTGCGAGACCTGCATATCTGGCTCTAT 3360  
Qy 1098 GlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAsnLeu 1117  
Db 3361 GGTGACCCGACATATGGAGGCAATGTGACCATGCAAGTGCATCAACGAGGACGATCCTG 3420  
Qy 1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137  
Db 3421 TGCAACACCAACACTGGCAAGTCTTCTGTACCAACAGGGGGTCAAGGAGAGGAGTGC 3480  
Qy 1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157  
Db 3481 CAGCTATGTGAGTGAATAATCGATACCGAGGAAACCCCTCTCAAGAGCAATGTACTAT 3540

Qy 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177  
Db 3541 ACCCTTCTCATTGACTATCAGTTTCACTTTTAGCTGTCCCGAAGACGACCGCTATTAC 3600  
Qy 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197  
Db 3601 ACAGCCATCAATTTTGGCTACTCTCTGATGAACAAACAGGATTTGGACATGTTTCATC 3660  
Qy 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217  
Db 3661 AATGCCCTCCAAAACCTTCAACCTCAACATCACTGGGCCACCTTTTCCAGCTGGAAAC 3720  
Qy 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237  
Db 3721 CAGACTGGAGAGAGGTCCTGTGTTTCAAAACCAACATCAAGGAATACAAAGATAGC 3780  
Qy 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257  
Db 3781 TTCTCTAATGAGAAATTTGATTTTGCACACCATCCAAACATTACTTCTTTTGTATGTC 3840  
Qy 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSer-----GlnHisSer 1274  
Db 3841 AGTAAATTTCATTGGCCCATCAAAATTCAGTAGAGTACTTCTTAACTGTCATTCG 3897  
RESULT 12  
AB093245 Mus musculus mRNA for mKIAA0548 protein. linear ROD 07-FEB-2003  
LOCUS  
DEFINITION  
ACCESSION AB093245  
VERSION AB093245.1 GI:26006172  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
REFERENCE  
AUTHORS Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Hara, Y., Nagase, T., Ohara, O. and Koga, H.  
TITLE Prediction of the coding sequences of mouse homologues of KIAA gene: I. The complete nucleotide sequences of 100 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 5693)  
AUTHORS Okazaki, N., Kikuno, R., Nagase, T., Ohara, O. and Koga, H.  
TITLE Direct Submission  
JOURNAL Submitted (04-OCT-2002) Hisashi Koga, Kazusa DNA Research Institute, Laboratory for Genome Informatics; 2-6-7 Kazusa-kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: mouse@kazusa.or.jp, Tel: 81-438-52-3919, Fax: 81-438-52-3918) The CREATE program supported by Japan science and technology corporation; cDNA full insert sequencing: Kazusa DNA Research Institute; cDNA library construction, clone selection and 5'- & 3'-end one pass sequencing.  
FEATURES  
Location/Qualifiers  
1..5693  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/clone="mbg03041"  
/tissue\_type="brain"  
/dev\_stage="adult"  
/note="vector:modified pBC SK-"  
1..5693  
/gene="mKIAA0548"  
1..3181  
/gene="mKIAA0548"  
/note="CDS is predicted by in silico analysis. Start codon is not identified."  
/codon\_start=2  
/evidence=not\_experimental  
/product="mKIAA0548 protein"

```

/protein_id="BAC41429.1"
/db_xref="GI:2606173"
/translation="AYDLTSRWLPLNHSVNVVVYVYHSHLALHDKKIYMGKIDST
GNVTNELRYPH:HNESWILLTPKADQVAVGSHAHITLASGRVLMVIFGHCPLYG
YISVVOEYDELENTWTSILHTOGALVQGGYGHSSVYDDRTKALYVHGKYKAFSANKYRL
ADDLVYDVDTQWMTILKDSRFRYLHTAVIVSGTMLVFGGNTHNDTSMHSAKCFSS
DMAVDIAADRASVLPRPELHDHVNRFHSAVLNSTWYVFGGNSILLSDLVFTLE
QCDARSEAACAVALPGPRICLDWTSQSRCTSMELATEBOAEKLSCEKSEKTLDDRC
DORTDCYCTANTNDCHNCDCVFNHSCTEGQISIAKIESCPKDNPMYVYCNKITS
RSCALDQNCWEPNQECIALPENICNGWHLVGNSSCLIKITITAKENYDNLAKSRNH
AFILASTIYKXDSNFSENEKFRPNHNTIFFVYNSFTWPIKIQIASFQSHFNDFVQFV
LIQWMPSESDAGFCOILBEPSTRGKLAATCINPLNGSCVCPERPANHSKQRTCALR
TACGECTSSSESCSNMKQCVDNAYVAFPPQCMWYTMSSCPENCSCYCTCS
HCLQPGCCWCTDPNNTGKCKIEGYPKVPKMSOASAGNVPPOPILNLSMCLDSR
YVNSRTHCPACCCNGHSHKINGQVCEKEDLTGKHCECTCISGFYGDPTNGSKQCPK
CNGHSLCINTGKFCITTKGVKDECCQICEVNRIOGNPLKGTCTYILLIDYQTF
LSQEDDRYTAINFATPDQNRDLDMFINASKNFLNLTWATSPAGTQDGEVPFV
SKTTIKYKXDSNFSENEKFRPNHNTIFFVYNSFTWPIKIQIASFQSHFNDFVQFV
TFPSFCLILLVAAVVMVKIKQSCWASRRREQLREMQQVWASRPFASVNVVALETDDEPP
DLIGGSIKTVKPIALPCPGKNAVLVSFVRLPRGLGIPPPGQGLAVASALVDIS
QQMPIVYKEKSGAVNRKQPPAQPGTCI"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 0 Length: 5683
Score: 5624.00 Matches: 1010
Percent Similarity: 97.73% Conservative: 25
Best Local Similarity: 95.37% Mismatches: 24
Query Match: 71.05% Indels: 0
DB: 10 Gaps: 0

```

US-09-787-097-12 (1-1429) x AB093245 (1-5683)

```

Qy 371 AlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValVal 390
Db 2 GCGTATGACTGACCTCTAGAGGATGGCTTCACCTAAACCATCTCTGTAACAGTGTGTT 61
Qy 391 ValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLys 410
Db 62 GTAAGATATGGTCATCTTTGGCATTTACATAAGGATAAAATCTACATGTATGGAGGAAA 121
Qy 411 IleAspProThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSer 430
Db 122 ATTGATTCACAGGAGGAACTGACCAATGAGCTGAGAGATATTTCAATTCATAATGAATCA 181
Qy 431 TrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHis 450
Db 182 TGGGTATTGTTAACTCCGAAGCTAAGGATCAGTATGCAGTGTGTTGGACACTCAGCACAC 241
Qy 451 IleValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeu 470
Db 242 ATTTGTTACACTGGCACTCGCCGCTGGGTCATGTTGGTCATCTTCGGTCATTTGCCCATC 301
Qy 471 TyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeu 490
Db 302 TATGGATATTAAGCGTTGTGAGGATATGACTTGGAAAGAACACATGAGATATATTA 361
Qy 491 HisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSerSerValTyrAspHisArg 510
Db 362 CATACTCAGGCTGCTCTTGTGCAAGGGGTTATGGCCACAGTAGTGTGTTTATGATGACAGG 421
Qy 511 ThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeu 530
Db 422 ACCAAGGCTCTGTAACCTTGTGCTGCTACAGGCTTTTCAGGCCCAACAAATACCGGCTT 481
Qy 531 AlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSer 550
Db 482 GCAGATGACTCTACAGATACGATGGGATCTCAGAGTGTGGACCATTCCTTAAGGACAGC 541
Qy 551 ArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGly 570
Db 542 CGATTTTCCGTTACTTGCATACAGCTGTGATGAGTGTGGAACCATGCTGTGTGTTTGA 601

```

```

Qy 571 GlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPhe 590
Db 602 GGGAAACACACACATGACATTCATGAGCCAGGTGCCAAATGCTTCTCTCGGACTTC 661
Qy 591 MetAlaTyrAspIleAlaCysAspArgTyrPheSerValLeuProArgProAspLeuHisHis 610
Db 662 ATGGCTTATGACATTTGCTTGTGACCGATGTCAGTGCTTCCAGACAGCTGAGCTCCATCAT 721
Qy 611 AspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGly 630
Db 722 GATGTCAACAGATTTGGCCATTGAGCAGCTTGTGTACAACAGCACCATTGATGTGTTCGGC 781
Qy 631 GlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAla 650
Db 782 GCCTTCAACAGCCTCTCTCTCAGTGAGCTTGTGCTTTTACCTTGGAGCAGTCGCGATGCA 841
Qy 651 HisArgSerGluAlaAlaCysLysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThr 670
Db 842 CACCGCAGTGAAGCTGCTTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACA 901
Qy 671 GlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLys 690
Db 902 CAGTCTGCTCGATGTACTCTCTGGAGTTGGCACTGAAGAACACAGCAAAAGTTAAAA 961
Qy 691 SerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCys 710
Db 962 TCAGAGTGTGTTTCTAAAGAACCTTGTGACCATGACAGATGTGACCCAGCACACAGATTGT 1021
Qy 711 TyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArg 730
Db 1022 TAGAGTGCACAGCCCAATACCAATGACTGCCACTGGTGTCAATGATCCTGTGTCCCTGTG 1081
Qy 731 AsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProLysAsp 750
Db 1082 AACCCAGCTGCACAGAGGCCAGATCTCCATGTCACAGTATGAGAGTTGCCCAAGGAT 1141
Qy 751 AsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsn 770
Db 1142 AACCCCATGTACTACTGCAATTAAGAAACACAGCTGCAGAGAGTGTCCCTTAGACCAAGAAC 1201
Qy 771 CysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProLysAsnIleCysGlyIle 790
Db 1202 TGCCAGTGGAGGCCCGGATCAAGAGTGCATGCCCTGCCGGAATATCTGTGGCAAT 1261
Qy 791 GlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAsp 810
Db 1262 GGCTGGCATTTGGTGTGAAACTGCTGTCTGAAATCACTACTGCTAAGGAGATATTATGAC 1321
Qy 811 AsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrGlnLys 830
Db 1322 AATGCTAAATGCTGTAGGAACCAACATGCCCTTTTGGCTTCCCTCAGATCCCAAGAG 1381
Qy 831 LysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLys 850
Db 1382 AAGGTGAGTTTGTCTTAAGCAGCTTCGATTAATGCAATCATCTCAAAAGTATGTCCAAG 1441
Qy 851 LeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGlu 870
Db 1442 CTCACCTGACTCCATGGGTGGTCTTCGGAAGATCAATGTGTCTTACTTGTGCTGGAG 1501
Qy 871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
Db 1502 GATATGTCTCCATTCACAAATAGTTTGTGTCAGTGGATGCCATCTGAGCCCATGTGCT 1561
Qy 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
Db 1562 GGCTTCTGTGGGATCTGTGACAGCCTAGTACTCGGGGATTAAGAGCTGCACCTGCATC 1621
Qy 911 AsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArg 930
Db 1622 AACCTCTCAATGGCAGCGTCTGTGAAAGCCCTGCAACACACAGTCCCAAGCAGTCCCG 1681
Qy 931 ThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMet 950

```



```
Db 1682 ACACCATGTGCCCTCGGACAGCGTGTGCGAGTGCACTAGCAGCAGCTCGAGTGCATG 1741
Qy 951 TrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPheProPhe 970
Db 1742 TGGTCAGCAACATGAAGCAGTGTGTGACTCCAAATGCTACGTAGCGCTCTCCCTTTT 1801
Qy 971 GlyGlnCysMetGluTyrTyrThrMetSerThrCysProGluAsnCysSerGlyTyr 990
Db 1802 GCCCAGTGTATGGAATGTATACGATGAGCAGCTGCCACCTGAAAAATGCTCTGGCTAT 1861
Qy 991 CysThrCysSerHisCysLeuGluGlnProGlyCysGlyTyrCysThrAspProSerAsn 1010
Db 1862 TGTAAGTGCAGCCTATGCTGGAGCAGCCAGCTGTGGTGTGCTACTGATCTAGCAAT 1921
Qy 1011 ThrGlyLysGlyCysIleGluGlySerTyrLysGlyProValLysMetProSerGln 1030
Db 1922 ACTGGAAAGGAAATGATATTGAGGCGAGCTATAAAGGACCTGTGAAGATCCCGTACAG 1981
Qy 1031 AlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAsp 1050
Db 1982 GCCTCTGAGGAAATGCTATCCACAGCCCTTCTGAACCTCAGCATGTGTCTAGAGGAC 2041
Qy 1051 SerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLys 1070
Db 2042 AGCAGATACAACCTGTCTTCACTTCACTGTCCAGCTTGCAGTCCAGCAATGTCTAGAGGAC 2101
Qy 1071 CysIleAsnGlnSerIleCysGlyLysCysGluAsnLeuThrGlyLysHisCysGlu 1090
Db 2102 TGCAATCAACAGAGTGTCTGAGAGTGTGAGGACCTGACCCAGGCAAGCATGTGCGAG 2161
Qy 1091 ThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLys 1110
Db 2162 ACCTGCATATCTGGCTTCTATGTGTGACCGCACTAATGGAGGCAAAATGTCCAGCATCAAG 2221
Qy 1111 CysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLys 1130
Db 2222 TGCAATGGCAGCGATCACTGTCAACACCAACACCGGCAAGTGTCTGTACACCAAAA 2281
Qy 1131 GlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnPro 1150
Db 2282 GTGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCGGAAATCTT 2341
Qy 1151 LeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSer 1170
Db 2342 CTCAAGGAACAATGCTACTATACCTCTCTCAATGATCACTACCTTACCTTTAGCCTGCC 2401
Qy 1171 GlnGluAspAspArgTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsn 1190
Db 2402 CAGGAAGACGCGCTACTACAGCCATCACTTTGTGGCTACTCTCTGATGACAAAAC 2461
Qy 1191 ArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAla 1210
Db 2462 AGGGAATTGGACATGTTCAATGCTTCAAAAACCTTCAACCTCAACATCACTGGGCGC 2521
Qy 1211 AlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsn 1230
Db 2522 ACCAGCTTCCAGCCGGAACCCAGACTGGAGAGAGGTCCTGTGTTCCTTCAAAAACCCAC 2581
Qy 1231 IleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsn 1250
Db 2582 ATCAGGAATACAAAGATAGCTTCTCTAATGAGAATTTGATTTTCGCAACCATCCAAAC 2641
Qy 1251 IleThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPhe 1270
Db 2642 ATCACTTCTTCTTTTATGTAGTACGTAATTTCACTTGGCCCATCAAAATTCAGATGCTTC 2701
Qy 1271 SerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPhe 1290
Db 2702 TCCAGCAGCAGCAACTTCATGAGACCTGGTACAGTCTTGTGATCTTCTTCAAGTTGTTTC 2761
Qy 1291 LeuSerLeuLeuValAlaAlaValTrpLysIleGlnSerCysTrpAlaSer 1310
```

```
Db 2762 CTCTCGCTGCTTCTGTGGCTGCAGTGGTCTGAAGATCAAGCAGAGCTGTTGGGCATCC 2821
Qy 1311 ArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSer 1330
Db 2822 AGCGGAGAGAGAACATTTCTCGGAGATGCAACAGATGCCAGCCGCCCTTGTCTTCT 2881
Qy 1331 ValAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyGlySerIleLys 1350
Db 2882 GTAAACGTTGCTTGGAAACAGATGAAGAACCTTCTCTGATCTCATTTGGGGAAGTATAAG 2941
Qy 1351 ThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSer 1370
Db 2942 ACCGTTCTTAAGCCCATGTCCTCGAGCCCTGCTTTGGTAACAAAGCCGAGTCTCTCT 3001
Qy 1371 ValPheValArgLeuProArgGlyLeuGlyGlyLeuGlyLeuProGlyGlnSerGlyLeu 1390
Db 3002 GTATTCGTGAGCTTCTCTCGAGACTGGGAGGAATCCCTCTCTCTGTCAGTCAGTCTC 3061
Qy 1391 AlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLys 1410
Db 3062 GCTGTGCCAGTGCCTGTGGACATTTCTCAGCAGATGCCAATAGTGTACAAGGAGAAG 3121
Qy 1411 SerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db 3122 TCAGGAGCTGTAAACAAACCGAAGACGACGCCCTGCACAGCTTGAACCTGCAT 3178

RESULT 13
AV442317 4140 bp mRNA linear PRI 15-NOV-2003
LOCUS Homo sapiens attractin-like protein mRNA, complete cds.
DEFINITION AV442317
ACCESSION AV442317
VERSION AV442317.1 GI:38230554
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4140)
AUTHORS Milton,N.G.N.
TITLE Attractin-like peptides
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4140)
AUTHORS Milton,N.G.N.
TITLE Direct Submission
JOURNAL Submitted (18-OCT-2003) R&D, NeuroDelta Ltd., 35 Hillsborough Road,
London SE22 8QE, UK
FEATURES
Location/Qualifiers
source 1..4140
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="10"
CDS 1..4140
/note="membrane receptor; similar to KIAA0534 protein"
/product="attractin-like protein"
/protein_id="AA814297.1"
/db_xref="GI:38230555"
/translaton="METGGRARTGTPQAPAGVWRAPAGGGGGGASSWLDGNSWLL
CYGLYALYAVQVSKPCRTGSCGRVNSTCLCDPGWVGDDQCHQGRFKLTFP
SGYLTDPGFINYKTKTLEGPVNAALRFNFHATECSWDHYVDGDSYAPLI
KGLTGLVPIRGNETVEVTGVALHFFSDAAYNLTGFINI FYGINSCPNCSGH
AKVTSVSVQVYCECDKYWKGEACDIPYKANCQSPDHGYODLTGKLCVNDSSWQ
GPDCLANVSTESVWLLPNVKPSPSVGRASHKAVLHGKEMWVIGGTYENYSFQWYL
NYNLESSINWVGTSRGLRYGHSALYQENIFMYGRIETNDGNVTDLWLVNTHS
QSWSTKTPVLVGHGQYAVEGSHAHINELSDRVWMLIIFGYSIAYGYTTSIOBYHIS
SNWLVPEVTEKGAIVQGGYHTSVDEITKSIYVHGKYKALPGNKYIGLVLDLYKEVNT
KTWILKESFARYLHSAVLINGAMLIFFGNTNDTSLNGACFSADFAYDIACDE
WKIPLKHLRDVNRFGHSAVINGSNMYIFGFSVLLNDILVYPNCKAFKFRDELC
KAGPGIKYVWKNKHSWESGNTNNILRAKCPKPTAASDDRCYRVACDCASTANTNG
COWCDDKCKISANSNCMSKYNKTKCHVNEQICNKLTKSCSKSLNLCWQDQDQC
QALPAHLGEGEWSHIGDACRVNSRENYNKLKLYCNLSGNLSLTKSEVFEVDE
IQKYTEQKRVSPWGLRKINISYMGWEDMSPTNTTLLQWLPFEPNDSPGFCAYLERAAVA
```





QY 630 yGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAl 650  
Db 1791 GGGATTTTCTAGTGACTCCTTAATGATATCCTTGTATACAGCCTCAAAATGCAAGGC 1850  
QY 650 aHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTirAsnTh 670  
Db 1851 TTTCCAGAGATGAAGAAGCTTTGTAATAATGCTGCTCCAGGATAAATGTGTTTGAAT-- 1908  
QY 670 rGlySerSerGlnCysIleSerTirAlaLeuAlaThrAspGluGlnGluLysLeuLy 690  
Db 1909 ----AAAATCACTGTGAATCTTGG-----GAATCTGGGAATCACTAATAATATCTTAG 1958  
QY 690 sSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCy 710  
Db 1959 AGCAAGTGCCTCTCTAAACAGCTCTTCTGATGACAGATGTTACAGATGACAGATTG 2018  
QY 710 sTyrSerCysThrAlaAsnThrAsnAspCysHisTirCysAsnAsp---HisCysValPr 729  
Db 2019 TGCCAGCTGTACTGCAATACAAATGCGTGCATGTTGATGACAGAAATGCATTTTC 2078  
QY 729 oArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProLy 749  
Db 2079 GGCAATAGTAATCTGAGT-----ATGCTGTCAAGAACTACACCAATGTCAATG 2129  
QY 749 sAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGl 769  
Db 2130 GAGAAATCAGCAG---ATTGTAAACAACTTACCAGCTGTAAAGCTGTTCACTAAACTT 2186  
QY 769 nAsnCysGlnTirGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGl 789  
Db 2187 GAATTCAGCAGTGGGATCAGACACAGAAATGCCAGCTTTTACAGCTCATCTTTGTGG 2246  
QY 789 yIleGlyTirHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTy 809  
Db 2247 AGAAGATGAGTATATTTGGGATGCTGTCTTAGAGTCAATTCAGTACAGAGAAACTA 2306  
QY 809 rAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrGl 829  
Db 2307 TGCAATGCAAACTTTATGCTATATCTTAAAGTGAATCTTGTCTTCAATTAACAACCTC 2366  
QY 829 nLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSe 849  
Db 2367 AAAAGAGTGAATTTGTTCTGGATGAATA-----CAGAAGTATATAC 2408  
QY 849 rLysLeuThrLeuThrProTirValGlyLeuArgLysIleAsnValSerTyrTirCysTir 869  
Db 2409 ACAACAGAAAGTATCACCTTGGGTAGGCTTGCGCAAGATCAATATATCTATTTGGGATG 2468  
QY 869 pGluAspMetSerProPheThrAsnSerLeuLeuGlnTirMetProSerGluProSerAs 889  
Db 2469 GGAAGACATGCTCTCTTTTACAAACACACACTACAGTGGCTTCTGGCAACCCCAATGA 2528  
QY 889 pAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCy 909  
Db 2529 TTTCTGGTTTGTGCATATCTGMAAGGGCTGCAGTGGCAGGCTTAAAGCTAAATCTCTTG 2588  
QY 909 sIleAsnProLeuAsnGlySerValCysGluArgProAla-----AsnHisSerAl 926  
Db 2589 TACATCTATGGCAATGGGCTTGTCTGTAAGAACCTGTTGTAGTCCAAATCAAAATGC 2648  
QY 926 aLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySe 946  
Db 2649 GAGCGCTGCMAAAGCCATGCTCTCTAGGACATCATGTTCCAACTGTACAGCAATGG 2708  
QY 946 rSerGluCysMetTirCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAl 966  
Db 2709 CATGGAGTGTATGTGTGGCAGCATACGAAACGATGTGTGACTCTAAATGCCATATATCAT 2768  
QY 966 aSerPheProPheGlyGlnCysMetGlnTirTyrThrMetSerThrCysProProGluAs 986  
Db 2769 CTCTTTTCCATATGGCAATGTCTAGAGTGGCAACT---GCCACTGTCTCCCTCAAAA 2825  
QY 986 nCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTirCysTh 1006

Db 2826 TTGTTCTGATTTGAGAACCTGTGCGACAGTGTTTTGGAAACAGCCTGGATGTGGTGTGCAA 2885  
QY 1006 rAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProVally 1026  
Db 2886 TGATCTTAGTAATACAGGAAGAGACATTCGATTTGAGGTTTTCACGGGGACCAATGAA 2945  
QY 1026 sMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMe 1046  
Db 2946 GCTT-----ATTGGAATGCACACAGTGTGATGATGATGATGATGATGATGATGAT 2990  
QY 1046 tCysLeuGluAspSerArgTyrAsnTirSerPheIleHisCysProAlaCysGlnCysAs 1066  
Db 2991 TTCCCCCAAGAAAGAACTATGAGTGGTCCCTTATCCAGTGTCCAGCTTGCAGGTGAA 3050  
QY 1066 nGlyHisSerLysCysIleAsnGlnSerIleCysGlyLysCysGluAsnLeuThrThrGl 1086  
Db 3051 TGGACATAGCATTGTCATCAATAATAATGTGTGCGAAACAGTGTAAATAATCTCACCAAG 3110  
QY 1086 yLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCy 1106  
Db 3111 AAAGCAGTGTCAAGATTGTATGCCAGTTATTTATGAGATCCCAACCAATGGTGGAGTG 3170  
QY 1106 sGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPh 1126  
Db 3171 CACAGCTGTACATGAGTGGCCATGCAATAATCTGTCTCATCTGCACACAGGAAATGTTT 3230  
QY 1126 eCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTy 1146  
Db 3231 CTGCACAACTTAAAGGAATAAAGGTGACCAATGCCAATTTATGTGACTCTGAAATGCTA 3290  
QY 1146 rGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheTh 1166  
Db 3291 TGTGGTAATCCACTTAGAGAAACATGTTATTAAGCCCTTTGATTTGATTTCAATTAC 3350  
QY 1166 rPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheValaIaThrPr 1186  
Db 3351 CTTTCAGCTTATTACAGGAAGATGATGCCACCATCTGCCTCAATAACTTTATAGCAAAACC 3410  
QY 1186 oAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAs 1206  
Db 3411 AGAAGCTCGAACAATACTGGATATATCAATTAATGATCAACAACTTTATCTCAA 3470  
QY 1206 nIleThrTirAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValVa 1226  
Db 3471 CATTACGTGTCTGTCCGTTCAACAGCTGGAACAATATCTGGGAGAGAGACTTCTATAGT 3530  
QY 1226 lSerLysThrAsnLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheAr 1246  
Db 3531 TTCCAGAATAATAAAGGAATACAGAGATAGTTTTTCTATGAAATAATTAACCTTAG 3590  
QY 1246 gAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTirProIleLysIl 1266  
Db 3591 AAGCAATCTTAACATTAATCTTATGTAGTACGTCAGCAACTTTCTCTGGCTATTAAAT 3650  
QY 1266 eGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPh 1286  
Db 3651 ACAGATTGCAATCTCAACAACAATAACAATCAATGAGACCTTGTGAGTGTGTTTGTGACCTT 3710  
QY 1286 ePheSerCysPheLeuSerLeuLeuValAlaAlaValTirPlyLysIleLysGlnSe 1306  
Db 3711 CTTCAGTGTGTTCTTCTATCTTATGCTGGTGGCTGCTGTGTATGGAAGATCAAAACAAAC 3770  
QY 1306 rCysTirAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetalaSerAr 1326  
Db 3771 TTGTTGGCTTCTCGAGGAGAGCAACTGCTTCGAGAACGACAGCAGATGGCCAGCG 3830  
QY 1326 gProPheAlaSerValAsnValAlaLeuGluThrAspGluProProAspLeuIleGl 1346  
Db 3831 TCCCTTTGCTTCTGTTGATGTAGTCTGGAAGTGGAGCTGAAACAAACAGAGTTTCTGCG 3890  
QY 1346 yGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysAl 1366



QY 313 TrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGlu 332  
DB 687 TGGCAAGGCTGCTGATGTTCTTGAATGTTCCCTCTACTGAGTCTTCTGAGTT----- 740  
QY 333 GlutyrSerAsnLeuLys-----LeuProArgAlaSerHisLysAlaVal 347  
DB 741 ---CTGCCAACGTTAAACCCCTTCAGTCTCTCTGTAGGTGGGCTTCACTAAAGCAGTT 797  
QY 348 ValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTyrAsn 367  
DB 798 TTACACGGGAATTTATGCTGGGTGATGCTGGATATCTTTAACTACAGTCTCTTTTCAA 857  
QY 368 MetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsn 387  
DB 858 ATGGTCTTAATTAACAATTTAGAAAGCAGTATATGG---AATGTAGGAACCTCCATCAAG 914  
QY 388 AsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyr 407  
DB 915 GGACCTCTCCAGAGATATGGACACTCTCTCTGTTTATATACAGAAACATCTTTATGTAT 974  
QY 408 GlyGlyLysIleAspProThr---GlyAsnValThrAsnGluLeuArgValPheHisIle 426  
DB 975 GGAGCGCAATTTGAACAAATATGCGCAATGTCACAGATGAATATGCGTTTTTAAACATA 1034  
QY 427 HisAsnGluSerTrpValLeuLeuThrProTyrAla-----LysGluGlnTyrAla 443  
DB 1035 CATAGTCAGTCATGAGTACAAAACCTCTACTGTTCTTTGGACATGGTCAGCAGTATGCT 1094  
QY 444 ValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuVal 463  
DB 1095 GTGGAGGGACATTTACAGCACAATATATGGAGTTGGATAGAGATGTTGTATGATCATTA 1154  
QY 464 IlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAsp 483  
DB 1155 ATATTGGATATCTGCATATATGTTATACAGCAGCATACAGGAATACCATATCTCA 1214  
QY 484 LysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHis 503  
DB 1215 TCAACACACTGGCTGTTCCAGAAACTAAAGAGCTATTGTTACAAAGGTGGATGGCCAT 1274  
QY 504 SerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPhe 523  
DB 1275 ACTAGTGTGTATGATGAATAATACAAAGTCCATTTATGTTATGAGGGGTATTAAGCAT 1334  
QY 524 SerAlaAsnLysTyrArgLeuAlaAspLeuTyrArgTyrAspValAspThrGlnMet 543  
DB 1335 CCAGGGAACAATATGCAATGCTGTTGATGATCTTTATAAATATCAAGTTAAACACTAAGACT 1394  
QY 544 TrpThrIleLeuLysAspSerArgPheArgTyrLeuHisThrAlaValIleValSer 563  
DB 1395 TGGACTATTTTGAAGAAAGTGGGTTGCCAGATACCTTCATTCAGCTGTTCTTATCAAT 1454  
QY 564 GlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAla 583  
DB 1455 GGAGCTATGCTTATTTTGGAGGAATACCAATATGACACTTCCTTTGAGTAACGGTGCA 1514  
QY 584 LysCysPheSerSerAspPheMetalTyrAspIleAlaCysAspArgTrpSerValLeu 603  
DB 1515 AAATGTTTTCTGCCGATTTCTGGCATATGACATAGCTTGTGTAATGGAATGGAATACTA 1574  
QY 604 ProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsn 623  
DB 1575 CCAAAACCAATTTCTATAGAGATGTCACAGATTTGGACACTCTCCAGTACTCATTAAC 1634  
QY 624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643  
DB 1635 GGGTCCATGTATATATTGGGGGATTTCTAGTGTACTCTTAATGATATCTCTGTATATAC 1694  
QY 644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaGlyProGly 663  
DB 1695 AAGCCTCCAAATTCGAAGGCTTTCAGAGATGAAGAACTTTGTAAAAATGCTGTGCTCCAGG 1754

QY 664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683  
DB 1755 ATAAATGTTGTTGGAAT-----AAAAATCACTGTGAATCTTGG-----GAATCTGGG 1802  
QY 684 GluGlnGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703  
DB 1803 AATACTAATAATAATTTCTAGCAAAAGTGCCTCTCTAAACAGCTCTTCTGTATGACAGA 1862  
QY 704 CysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723  
DB 1863 TGTTACAGATATGACAGATTGTCAGCTGTACTGCCAATACAAATGGGTGCCAATGGTGT 1922  
QY 724 AsnAsp---HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePhe 742  
DB 1923 GATGACAAAGAAATTCGCAATTCGGCAATAGTAATCTGCAGT-----ATGCTGTCAAG 1973  
QY 743 ArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCys 762  
DB 1974 AACTACACCAAAATGTCATGTGAGAAATGAGCAG---ATTGTATAACAACTTACCAGCTGT 2030  
QY 763 ArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGlyCysIleAla 782  
DB 2031 AAAAGCTGTTTCACTAAACTTGAATTCGCAATTCGAGTGGGATCAGAGACAGCAAGAAATCCAGGCT 2090  
QY 783 LeuProGluAsnIleCysGlyIleGlyTyrPheIleValGlyAsnSerCysLeuLysIle 802  
DB 2091 TTACAGCTCATCTTGTGGAGAGATGGAGTCATATTTGGGATGCTTGTCTTAGAGTC 2150  
QY 803 ThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeu 822  
DB 2151 AATTCAGTAGAGAAACTATGACAAATGCAAAACTTTATTGCTTATATCTTACTGGAAAT 2210  
QY 823 LeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMet 842  
DB 2211 CTTCCTTCATTACCACTCAAAAGAGTAGAATTTGTTCTGATGAATA----- 2261  
QY 843 GluSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIle 862  
DB 2262 -----CAGAAGTATACAAACAGAAAGTATACCTTTGGTAGGCTTGGCAAGATC 2312  
QY 863 AsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrp 882  
DB 2313 AATATATCTTATTTGGGATGGAGACATGTCTCTTTTACAAACACAACTACATGCG 2372  
QY 883 MetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArg 902  
DB 2373 CTTCCTGGCAACCAATGATCTGGGTTTGTGCATATCTGGAAAGGCTGCGATGGCA 2432  
QY 903 GlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAla 922  
DB 2433 GGCTTAAAGCTAATCTCTGTACATCTATGGCAAAATGGCCTTGTCTGTGAAAACTGTT 2492  
QY 923 -----AsnHisSerAlaLysGlnCysArgThrProCysAlaLeuAlaGlyThrAlaCys 939  
DB 2493 GTTAGTCCAAATCAAAATGCGAGCCGTGCAAAAGCATGCTCTCTGAGGACATCATGT 2552  
QY 940 GlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959  
DB 2553 TCCAACTGTACAAGCAATGGCATGGATGTATGTGTGTGTCAGCAGTACGAAACGATGTT 2612  
QY 960 AspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMet 979  
DB 2613 GACTTATGCTATATCATCTCTTTCCATATGACAAATGCTTAGAGTGGCAACT--- 2669  
QY 980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999  
DB 2670 GCCACTGCTCCCTCAAAATTTCTGGATTGAGAACTCTGGACAGTGTGTGGAAACAG 2729  
QY 1000 ProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGly 1019  
DB 2730 CCTGGATGCTGCTGTGCAATGATCTAGTAATACAGGAAGAGGACATTCATTGAAGGT 2789  
QY 1020 SerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGln 1039

Db 2790 TCITCAGGGGACCAATGAAGCTT-----ATTGGAATGCACCAATAGAG 2834  
Qy 1040 ProLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheLeuHis 1059  
Db 2835 ATGGTTCTTGACACCAATCTTTGCCCAAGAAAGAACTATGAGTGTCTTATCCAG 2894  
Qy 1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerLysCysGluLys 1079  
Db 2895 TGTCCAGCTTGCAGGTGAATAGACATAGCACTTGCATCAATAATAATGTGTGCGAAGCAG 2954  
Qy 1080 CysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAsp 1099  
Db 2955 TGTAAATAATCTACACAGAAAGCAGTGTCAAGATGTATGTCAGCATTTATTTAGGAT 3014  
Qy 1100 ProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsn 1119  
Db 3015 CCAACCAATGCTGACAGTGCACAGCTTGTACATGCACTGGCCATGCAATATCTGTCTAT 3074  
Qy 1120 ThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeu 1139  
Db 3075 CTGCACACAGAAATGTTCTGCACAACTAAAGGAATAAAGGTGACCAATGCCAATTA 3134  
Qy 1140 CysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeu 1159  
Db 3135 TGTGACTCTGAAATCGCTATGTTGTTGATCCACTTAGAGGAACATGTTATTACAGCTT 3194  
Qy 1160 LeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrThrAla 1179  
Db 3195 TTCAATTGATTATCAATTTACCTTCAGCTTATTACAGGAAGATGATCGCCACCATATGCC 3254  
Qy 1180 IleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAla 1199  
Db 3255 ATAACTTTATACCAACCCAGACAGTGCACAAATAATCTGGATATATCAATTAATGCA 3314  
Qy 1200 SerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAla 1219  
Db 3315 TCAACAACACTTTAATCTCAACATACCTGCTGTCTGCTTCAACAGCTGGACAATATCT 3374  
Qy 1220 GlyGluLysMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSer 1239  
Db 3375 GGGGAAGAGACTTCTATGTTTCCAAAGATAATATAAGGAATACAGAGATAGTTTTC 3434  
Qy 1240 AsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsn 1259  
Db 3435 TATGAAAATTTAACTTTAGAACCAATCTTAACATTACGTTCTATGTGTACGTACAGAAC 3494  
Qy 1260 PheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeu 1279  
Db 3495 TTTTCCTGCGCTATTAAATAACAGATTGCAATCTCACAAACAATACAAATCATGAGACCTT 3554  
Qy 1280 ValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaVal 1299  
Db 3555 GTGCACTTTTGTACCTTCTTCAGTTGTTCTTCTATTCCTATTCCTGCTGTGTGTGTG 3614  
Qy 1300 ValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGlu 1319  
Db 3615 GTATGGAAGATCAAAACAATCTTTGGGCTTCTGACGAGAGAGACAACTTCTTCGAGAA 3674  
Qy 1320 MetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGlu 1339  
Db 3675 CACAGCAGATGCGCAGCGTCCCTTTGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3734  
Qy 1340 GluProAspLeuIleGlySerIleLysThrValProLysProIleAlaLeuGlu 1359  
Db 3735 GAAACAAACAGAGTTCTCTGAGGGCCCAATAGAGGGGGCCACCAAGCCCAATGCCATTGAA 3794  
Qy 1360 ProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeu 1379  
Db 3795 CCATGTGCGGAACAGAGCTCTGTTCTGATGTTCTTTGTTCTTACACAGAGGATCA 3854  
Qy 1380 GlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIle 1399

Db 3855 TCAGTGCCTCCCTCCCTGGGCGAGTCAGGCCCTTCAATTGCCAGTGCCTTAATAGATATT 3914  
Qy 1400 SerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGln 1419  
Db 3915 TCACAAAGAAAGCTTCAGATAGTAAAGATAAGACTTCTGGAGTCCGGAATCGAAACAC 3974  
Qy 1420 GlnProAlaGlnProGlyThrCysIle 1429  
Db 3975 CTTTCAACACGCTCAA---GGAACTTGTGTC 4001  
RESULT 15  
ARI164818  
LOCUS ARI164818 2625 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 18 from patent US 6274339.  
ACCESSION ARI164818  
VERSION ARI164818.1 GI:16238029  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2625)  
AUTHORS Moore,K. and Nagle,D.Lynn.  
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity  
JOURNAL Patent: US 6274339-A 18 14-AUG-2001;  
FEATURES location/Qualifiers  
source 1. 2625  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Zred. No.: 6.29e-250 Length: 2625  
Score: 4136.50 Matches: 759  
Percent Similarity: 95.48% Conservative: 2  
Best Local Similarity: 95.23% Mismatches: 22  
Query Match: 52.26% Indels: 14  
DB: Gaps: 3  
US-09-787-097-12 (1-1429) x ARI164818 (1-2625)  
Qy 2 ValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 21  
Db 10 GTGCGCGCAGCGCGCACTGAGGCAAGCTCAGGAGGAGGAGCGCGCGCGCGCGCGCG 69  
Qy 22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41  
Db 70 CTGCGGCGCAGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 129  
Qy 42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56  
Db 130 GGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 189  
Qy 57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 76  
Db 190 CTGCTGCGCG---CTGCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 225  
Qy 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 96  
Db 226 CTGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 285  
Qy 97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116  
Db 286 GAGGCCCAAGGAATGTGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 345  
Qy 117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136  
Db 346 CAGTGGCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCT 405  
Qy 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156  
Db 406 AGACTAACTGGATCTTCTGGGTTTGTGACAGATGACCTGGAAATATATAAATACAAACG 465

QY 157 LysCysThrTrpLeuIleGluGlnProAsnArgIleMetArgLeuAArgPheAsnHis 176  
DB 466 AAGTGCAGTGGCTCAITTAAGAGCAGCCAAATAAGATAGACTTCGTTTCAATCAT 525  
QY 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196  
DB 526 TTGCTCAGAGTGTAGTGTGGACCAATTATATGTTATGATGGGACTCAATTATATGCA 585  
QY 197 ProLeuValAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216  
DB 586 CCGCTAGTTGCTGCAATTTAGTGGCTCATTTCTCTGAGAGAGATGGCAATGAGACTGTC 645  
QY 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236  
DB 646 CCGTAGGTTGTTGCCACATCAGTTATGCCCTGCTGCAATTTTTTAGTATGCTGCTTAT 705  
QY 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256  
DB 706 AATTTCAGCTGATTTAATATTACTTTACAGTTTGTATATGTTGCCAAATAAATGCTGCAGC 765  
QY 257 ArgGlyGluCysLeuIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276  
DB 766 CGAGAGAGTGTAAAGATCAGTAATAGCAGCGATACCTGTTGAATGGAATGTTCTGAAGAC 825  
QY 277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296  
DB 826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTCTCATCGA 885  
QY 297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316  
DB 886 GGCATCTGCANATCAAGTATGTCTCAGAGATGCTCTCTCTCAGACTGCAGGGTCTCT 945  
QY 317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336  
DB 946 GGATGTTTCAGTTCCCTGTCACCACTAACCGTCAATTTGGACTCGAGAGGAATAATCTAAC 1005  
QY 337 LeuLysLeuProArgAlaSerHisLysAlaValAlaValAsnGlyAsnIleMetTrpValVal 356  
DB 1006 TTAAGCTCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGGAAACATTTATGTGGTTGTT 1065  
QY 357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376  
DB 1066 GGAGGATATATGTTCAACACCTCAGATATATACATGGTTCTAGCGTATGACCTTCTCTCT 1125  
QY 377 ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396  
DB 1126 AGGAGTGGCTCCACTAAACCGTCTGTGAACAAATGTTGTTAGATATGCTCATTTCT 1185  
QY 397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyLysIleAspProThrGlyAsn 416  
DB 1186 TTGGCATTTATACAAGGATATAATTTACATGTATGGAGGAAATAATGATTCAACTGGGAAT 1245  
QY 417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro 436  
DB 1246 GTGACCAATGAGTTGAGAGTTTTCATTTCAATTCATATGAGTATGGGTGTTGTGACCCCT 1305  
QY 437 LysAlaLysGluGlnTyrAlaValAlaValGlyHisSerAlaHisIleValThrLeuLysAsn 456  
DB 1306 AAGGCAAAAGAGCAGTATGCGTGGTGGGCACTCTGCACACATTTGTTACACTGAAGAT 1365  
QY 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476  
DB 1366 GGCCGAGTGGTGTATGCTGGTCATCTTTGGTCACTGCCCTCTCTATGATATATTAAGCAAT 1425  
QY 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496  
DB 1426 GTGACGAATATGATTTGGATAGAACACATGGAGTATATATACACACCCAGGGTGGCCTT 1485  
QY 497 ValGlnGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516  
DB 1486 GTGCAAGGGGTTACGGCCATAGCAGTGTGTTTACGACCATAGGACCCAGGGCCCTTATACGTT 1545  
QY 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536

DB 1546 CATGTGGCTACAGGCTTTTCAGTGCATAAGTACCGCTTTCAGATGATCTCTACCGA 1605  
QY 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeu 556  
DB 1606 TATGATGTGGATACCCAGATGTGGACCATTTCTTAGGACAGCCGATTTTTCGTTACTTG 1665  
QY 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576  
DB 1666 CACACAGCTGTGATAGTGGGAACCATGCTGGTGTGTTGGGGGAAACACACACAATGAC 1725  
QY 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596  
DB 1726 ACATCTATGAGCCATGCGCCAAATGCTTCTCTTCAGATTTTCATGGCCCTATGACATGCC 1785  
QY 597 CysAspArgTrpSerValLeuProAspProAspLeuHisHisAspValAsnArgPheGly 616  
DB 1786 TGTGACCGCTGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATCTCAACAGATTGGC 1845  
QY 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
DB 1846 CATTCAGCAGCTTTACACAACAGCAGCAATGATGTGTCGGTGGTTTCAATAGTCTCTCTC 1905  
QY 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
DB 1906 CTGAGGACATCTGCTGATTTCACTCGAACAGTGTGATCGCATTCGGAGTGAAGCGCT 1965  
QY 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676  
DB 1966 TGTTAGCAGCAGGACCTGGTATTCGTTGTTGTCGAACACAGGGTCTCTCAGTGTATC 2025  
QY 677 SerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696  
DB 2026 TCGTGGCGCTGGCACTGATGAACAGAAAGTTAAATCAGAAATGTTTTTCCAAA 2085  
QY 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
DB 2086 AGAATCTTTGACCATGACATGTGACACACACAGATTTGTTACAGCTGCACAGCCAC 2145  
QY 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
DB 2146 ACCAATGACTGCCACTGGTGCATGACCATTTGTTGCCAGGAACCAACAGCTGCTCAGAA 2205  
QY 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
DB 2206 GGCCAGATCTCCATTTTAGTATGAGATTTGCCCAAGGATTAACCTATGACTACTGT 2265  
QY 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776  
DB 2266 AACAAAGAGACCACTGCAGGAGCTGTGCCCTGGACCCAGAACTGCCAGTGGAGCCCGG 2325  
QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCys-GlyIleGlyTrp 792  
DB 2326 AATCAGGAGTGCATTTGCCCTGCCCGTAGGCTTTGCAGGGTTCATCTTGG 2374

Search completed: March 1, 2004, 05:12:59  
Job time : 11835 secs

***This Page Blank (uspto)***



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: March 1, 2004, 00:45:20 ; Search time 6892 Seconds  
(without alignments)  
6191.680 Million cell updates/sec

Title: US-09-787-097-12  
Perfect score: 7915  
Sequence: 1 MVAAAAPEARLRRRTAATA.....KSGAVRNKQPPAQPQTCTI 1429

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues  
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.spool -DEV=xlp  
-O=/cgn2\_1/USPTO.spool\_p/US09787097/runat\_26022004\_062935\_23428/app\_query.fasta\_1.1607  
-DB=EST\_QPMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0 -LIST=45  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -MODE=LOCAL  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US09787097 @cgn 1.1 4337 @runat\_26022004\_062935\_23428 -NCPU=6 -ICPU=3  
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :  
EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: em\_gss\_hum:\*  
18: em\_gss\_inv:\*  
19: em\_gss\_pln:\*  
20: em\_gss\_vrt:\*  
21: em\_gss\_fun:\*  
22: em\_gss\_mam:\*  
23: em\_gss\_mus:\*  
24: em\_gss\_pro:\*  
25: em\_gss\_rod:\*  
26: em\_gss\_pbg:\*  
27: em\_gss\_vrl:\*  
28: gb\_gss1:\*

29: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6217	78.5	3976	29	AY418587 Homo sapi
2	5980	75.6	3976	29	AY418589 Mus muscu
3	5018	63.4	3671	29	AY418588 Pan trogl
4	1746.5	22.1	1201	13	BX440935 BX440935
5	1619	20.5	964	13	BU506373 AGENCOURT
6	1491	18.8	840	14	CB520374 UI-M-GIO-
7	1421.5	18.0	933	12	BG678679
8	1402	17.7	1545	11	AK050882 Mus muscu
9	1397	17.7	770	13	BY761276 BY761276
10	1350	17.1	735	13	BY761390 BY761390
11	1315	16.6	777	14	CF745736 UI-M-GVO-
12	1310	16.6	681	12	BM783739 K-EST0061
13	1305	16.5	680	12	BM783258 K-EST0061
14	1305	16.5	680	12	BM783773 K-EST0061
15	1296	16.4	683	12	BM783788 K-EST0061
16	1296	16.4	677	12	BM783718 K-EST0061
17	1292	16.3	673	12	BM783685 K-EST0061
18	1291	16.3	1836	11	BC031517 Homo sapi
19	1290	16.3	677	12	BM783821 K-EST0061
20	1254	15.8	722	13	EX669806 EX669806
21	1233	15.6	792	14	CD844584 AGENCOURT
22	1233	15.6	962	13	EX431593 EX431593
23	1231.5	15.6	714	10	BE650220 UI-M-BH2.
24	1226	15.5	674	12	BG563729 602584512
25	1225	15.5	917	10	BF672370 602150605
26	1208.5	15.3	940	13	BQ956223 AGENCOURT
27	1206	15.2	632	12	BM773003 K-EST0057
28	1206	15.2	632	12	BM783264 K-EST0057
29	1191	15.0	716	14	CF745901 UI-M-GVO-
30	1191	15.0	739	9	AL629411 AL629411
31	1187	15.0	706	12	BI089332 602853292
32	1183	14.9	809	13	BU286532 603605969
33	1179	14.9	663	14	CF745835 UI-M-GVO-
34	1163	14.7	743	14	CB519923 UI-M-GIO-
35	1161	14.7	633	10	BE894518 601433126
36	1130	14.3	626	10	BE980188 UI-M-BG2-
37	1126	14.2	953	10	BF313142 601897661
38	1125.5	14.2	1734	11	BC029592 Homo sapi
39	1125	14.2	591	12	BI850289 imageqc.1
40	1115.5	14.1	902	13	BU151781 AGENCOURT
41	1112	14.0	710	13	BU377616 603811614
42	1110	14.0	991	10	BE880905 601430448
43	1108	14.0	780	13	BU378893 603813635
44	1088	13.7	653	14	CB446934 700912 MA
45	1079	13.6	641	10	BB664616 BB664616

ALIGNMENTS

RESULT 1  
AY418587  
LOCUS  
DEFINITION Homo sapiens ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,  
3976 bp DNA linear GSS 17-DEC-2003  
ACCESSION AY418587  
VERSION AY418587.1 GI:39774547  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 3976)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,R., Thomas,P., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.

TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)

PUBMED 14671302

REFERENCE 2 (bases 1 to 3976)

AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,R., Kejarival,A., Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., Adams,M.D. and Cargill,M.

TITLE Direct Submission

JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA

COMMENT These sequences were made by sequencing genomic exons and ordering them based on alignment.

FEATURES

source 1..3976  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
<1..>3976  
/gene="ATRN"  
/locus\_tag="HCM6607"

ORIGIN

Alignment Scores:  
Pred. No.: 0 Length: 3976  
Score: 6217.00 Matches: 1139  
Percent Similarity: 86.10% Conservative: 1  
Best Local Similarity: 86.03% Mismatches: 184  
Query Match: 78.55% Indels: 0  
DB: 29 Gaps: 0

US-09-787-097-12 (1-1429) x AY418587 (1-3976)

QY 106 ValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpVal 125  
DB 2 GTCAACGGCGGTGCTGCACACCTGGCAGCGCCAGTGGCTGCCCCGGCTGGGTG 61

QY 126 GlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheVal 145  
DB 62 GCGCAGCAATGCCAGCACTCGCGGGCGCGCTTCAGNNNNNNNNNNNNNNNNNNNNNN 121

QY 146 ThrAspGlyProGlyAsnTrpLysTrpLysCysThrTrpLeuIleGluGlyGln 165  
DB 122 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 181

QY 166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185  
DB 182 CCAAAATAGAATAATGAGACTTCGTTCAATCATTTTGTCTACAGAGTGTAGTTGGACCAT 241

QY 186 LeuTrpValTrpAspGlyAspSerIleTrpAlaProLeuValAlaAlaPheSerGlyLeu 205  
DB 242 TTATATGTTTATGATGGGACCTCAATTTATGCACCGCTAGTTGCTGCAATTTAGNNNNNN 301

QY 206 IleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyr 225  
DB 302 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 361

QY 226 AlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyr 245  
DB 362 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 421

QY 246 SerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSer 265  
DB 422 NNTTTTGTATGTTCCCAATTAATGCTCAGGCCCGGAGAGTGTAAAGATCAGTAATAGC 481

QY 266 SerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIlePro 285  
DB 482 AGCGATACTGTTGAATGTGAATGTTCTGAAACTCGAAAGGTTGAAGCATGTGACATTCCT 541

QY 286 HisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArg 305  
DB 542 CACTGTACAGACAACCTGGTGTTCCTCATCGAGGCATCTGCAATTCAAGTGTATGCAGA 601

QY 306 GlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsn 325  
DB 602 GGATGCTTCCTGCTTCTCAGACTCGCAGGGTCTGGATGTTTCAGTTCTGTACAGGTAAC 661

QY 326 GlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLys 345  
DB 662 CAGTCACTTTGGACTCGAGAGGAATATTTAACTTAAAGCTCCCGAGAGCATCTCATAAA 721

QY 346 AlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAsp 365  
DB 722 GCTGTGTCATCAATGGAACATTATGTGGTGTGTGGAGGATATATGTTCAACCACTCAGAT 781

QY 366 TyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSer 385  
DB 782 TATAACATGCTCTAGCGTATGACCTTCTTAGGGAGTGGCTTCCACTAAACCGTTCT 841

QY 386 ValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyr 405  
DB 842 GTGAACAATGCTGTTGTAGATATGTCATTCTTTGGCATTATACAGNNNNNNNNNNNN 901

QY 406 MetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGluLeuArgValPheHis 425  
DB 902 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 961

QY 426 IleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValVal 445  
DB 962 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1021

QY 446 GlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePhe 465  
DB 1022 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1081

QY 466 GlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsn 485  
DB 1082 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1141

QY 486 ThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSer 505  
DB 1142 ACATGGAGTATATTACACCCAGGGTGGCTTGTGCAAGGGGGTTACGGCCATACAGT 1201

QY 506 ValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAla 525  
DB 1202 GTTTACGACCATAGGACCGGGCCCTATAGTTTCATGTTGGCTACAAGGCTTTCAGTGCC 1261

QY 526 AsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThr 545  
DB 1262 AATAAGTACCGGTTCCAGATGATCTTACCGATATGATGTGGATACCCAGATGTGGACC 1321

QY 546 IleLeuLysAspSerArgPhePheAspTyrLeuHisThrAlaValIleValSerGlyThr 565  
DB 1322 ATTCTTAAGACACCGCATTTTTCGTTACTTGCACACAGCTGTGATGAGTGAACCC 1381

QY 566 MetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCys 585  
DB 1382 ATGCTGCTGTTGGAGAAACACACAATGACATCTATGAGCCATGCGCCCAATGC 1441

QY 586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArg 605  
DB 1442 TTCTCTTCAGATTTCATGGCCTATGACATTTGNNNNNNNNNNNNNNNNNNNNNNNNNN 1501

QY 606 ProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThr 625  
DB 1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1561

QY 626 MetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSer 645  
DB 1562 ATGTATGTGTGTTGGTGGTTTCAATAGTCTCTCTCTCAGCGACATCTGTGTTACCTCG 1621

QY 646 GluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArg 665  
DB 1622 GAACAGTGTGATCGCATCGAGTGAAGCCGCTGTTTACGACGAGGACCTGGTATTCGG 1681  
QY 666 CysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGln 685  
DB 1692 TGTGTGTGAACACAGGGTGTCTCAGTGTATCTGTGGCGCTGGCACTGATGAACAA 1741  
QY 686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705  
DB 1742 GAAGAAAAGCTTAAATCAGATGTTTCCAAAGAACTCTTGACCATGACAGATGTGAC 1801  
QY 706 GlnHisThrAspCysTyrosSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAsp 725  
DB 1802 CAGCACACAGATTTGTACAGCTGCACAGCAACCAATGATGCTGCCCTGGTGCATGAC 1861  
QY 726 HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrglu 745  
DB 1862 CATTGTGTCCCGAGAACACACAGCTGCTCAGAAGGCCAGATCTCCATTTTATGATGAG 1921  
QY 746 AsnCysProLysAspAsnProMetTyrosCysAsnLysLysThrSerCysArgSerCys 765  
DB 1922 AATTGCCCCACAGATAAACCCTATGCTACTGTATAAAGAACACAGCTGCAGGAGCTGT 1981  
QY 766 AlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGlu 785  
DB 1982 GCCCTGGACAGAACTGCCAGTGGAGAGCCCGGAATCAGAGTGCATTTGCCCTGCCGAA 2041  
QY 786 AsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAla 805  
DB 2042 AATACTGTGGCATTTGGCTGGCATTTGGTTGGAACTCATGTTTGAAATTAATACTACTGCC 2101  
QY 806 LysGluAsnTyrosAspAsnAlaLysIlePheCysArgAsnHisAsnAlaLeuAlaSer 825  
DB 2102 AAGGAGAAATATGACAATGCTAAATGTTCTGTAGGAACCAATGCCCTTTTGGCTTCT 2161  
QY 826 LeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSer 845  
DB 2162 CTTACAAACCAGAGAGTAGAATTTGTCTTTAGCAGCTGCGAATTAATGCAGTCACT 2221  
QY 846 GlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSer 865  
DB 2222 CAGAGCATGTCCAAAGCTCACTTAACCCATGGTGGCTTTCGGAAGATCAATGTGTCC 2281  
QY 866 TyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSer 885  
DB 2282 TACTGTGTGGAGATATGTCCTTACAAATAGTTTACTACAGTGGATGCCGCT 2341  
QY 886 GluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLys 905  
DB 2342 GAGCCAGTGTGCTGGATTTCTGTGGAATTTATCAGAACCAGTACTCGGGAGCTGAAG 2401  
QY 906 AlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSer 925  
DB 2402 GTTGCAACTGTGATCAACCACTCAATGTGTGTCTGTGAAGGCCCTGCAACCCACAGT 2461  
QY 926 AlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGly 945  
DB 2462 GCTAACAGTGGCGACACCATGTGCTTGAGACAGCATGTGGAGATTGCACACGCGC 2521  
QY 946 SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrosVal 965  
DB 2522 AGCTCTGTAGTGTGTGTGTCAGCAACATGAAGCAGTGTGTGCACTCCAATGCCATGTG 2581  
QY 966 AlaSerPheProPheGlyGlnCysMetGluTrpTyrosThrMetSerThrCysProProGlu 985  
DB 2582 GCTCTCTCCCTTTGGCCAGTGTATGGAATGGTATACATGAGCAGCCTGCCCTCCCTGAA 2641  
QY 986 AsnCysSerGlyTyrosCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCys 1005  
DB 2642 AATTGTTACGGCTACTGTATCTGATCTGATCTGAGCAACCAACAGGCTGTGGCTGTGT 2701  
QY 1006 ThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrosGlyProVal 1025

DB 2702 ACTGATCCCGACGAACTACTGGCAAGGGAAATGATAGAGGTTCTCTATAAGGACCGAGT 2761  
QY 1026 LysMetProSerGlnAlaProThrGlyAsnPheTyrosProGlnProLeuLeuAsnSerSer 1045  
DB 2762 AAGATGCTTCGCAAGCCCTCAGAGAAATTTATCCACAGCCCTGCTCAATTCACG 2821  
QY 1046 MetCysLeuGluAspSerArgTyrosAsnTrpSerPheIleHisCysProAlaCysGlnCys 1065  
DB 2822 ATGTGTCTAGAGGACAGCAGATCAACTGGTCTTTTATTCTACCTGTCTCCAGCTTCCCAATGC 2881  
QY 1066 AsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThr 1085  
DB 2882 AACGCCACAGTAATCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTTGACCACA 2941  
QY 1086 GlyLysHisCysGluThrCysIleSerGlyPheTyrosGlyAspProThrAsnGlyGlyLys 1105  
DB 2942 GGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGCATCCCAATGGAGGAAA 3001  
QY 1106 CysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCys 1125  
DB 3002 TGTGAGCATGCAAGTGCAATGGGACGCTCTCTGTGCAACCAACACGCGGCAAGTGC 3061  
QY 1126 PheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArg 1145  
DB 3062 TTCTGACACCAAGGCGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAATCGA 3121  
QY 1146 TyrGlnGlyAsnProLeuArgGlyThrCysTyrosTyrosThrLeuLeuLeuAspTyrosGlnPhe 1165  
DB 3122 TACCAGGAAACCCCTCTCAGAGGAACATGTTATTATCTCTTCTTATTGACTATCAGTTC 3181  
QY 1166 ThrPheSerLeuSerGlnGluAspAspArgTyrosThrAlaIleAsnPheValAlaThr 1185  
DB 3182 ACCTTTAGTCTATCCCGAGGAAGATGATCGCTATTACACGCTATCAATTTTGGCTACT 3241  
QY 1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205  
DB 3242 CCTGACGAACAAACAGGATTTGGACATCTTCATCAATGCCTCCAAGAAATTTCAACCTC 3301  
QY 1206 AsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal 1225  
DB 3302 AACATACCTCGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAAGATGCCCTGTT 3361  
QY 1226 ValSerLysThrAsnIleLysGluTyrosLysAspSerPheSerAsnGluLysPheAspPhe 1245  
DB 3362 GTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTT 3421  
QY 1246 ArgAsnHisProAsnIleThrPhePheValTyrosValSerAsnPheThrTrpProIleLys 1265  
DB 3422 CGCAACCAACCAAAATACACTTTTCTTTTATGTCTAGTAAATTTCACTCGCCCATCAAA 3481  
QY 1266 IleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThr 1285  
DB 3482 ATTCAAGATTGCTTCTCTCAGCACAGCAATTTATGGACCTGGTACAGTTCCTTCGTGACT 3541  
QY 1286 PhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrpLysIleLysGln 1305  
DB 3542 TTTCTCAGTTGTTTCTCTCTTGTCTCCTGTGTGCTGTCTGTGTTTGAAGATCAAAACA 3601  
QY 1306 SerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSer 1325  
DB 3602 AGTTGTTGGGCGCTCCAGACCTAGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGC 3661  
QY 1326 ArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIle 1345  
DB 3662 CGTCCCTTTGCCCTGTAAATGTCCGCTTGAAACAGATGAGGAGCCTCTCTGATCTTATT 3721  
QY 1346 GlyGlySerLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys 1365  
DB 3722 GGGGGAGTATAAGACTGTTCCTCAACCCCAATTCGACTGGAGCGCTGTTTGGCAACAAA 3781  
QY 1366 AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProPro 1385

```

Db      3782 GCGCGTGTCTCTCTCTGTGTGTGGCTCCCTCGAGGCTCGGTGGCATCCCTCTCTCT 3841
Qy      1386 GlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIle 1405
Db      3842 GGGCAGTCAGGCTTCTGTGTGGCAGCGCCCTGTGTGGACATTTCTCAGCAGATGCCGATA 3901
Qy      1406 ValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnPro 1425
Db      3902 GTGTACAAGGAGAGTTCAGAGCGGTGAGAAACCGAGCAGCAGCCCTCGACAGCT 3961
Qy      1426 GlyThrCysIle 1429
Db      3962 GGGACCTGCATC 3973

RESULT 2
AY418589
LOCUS      3976 bp      DNA      linear      GSS 17-DEC-2003
DEFINITION Mus musculus ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,
GENOMIC SURVEY SEQUENCE.
ACCESSION AY418589
VERSION   AY418589.1 GI:39774549
KEYWORDS  GSS.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3976)
            Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
            Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
            Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Inferring nonneutral evolution from human-chimp-mouse orthologous
            gene trios
            Science 302 (5652), 1960-1963 (2003)
            14671302
REFERENCE 2 (bases 1 to 3976)
            Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
            Todd,M.A., Tanenbaum,D.M., Civallo,D.R., Lu,F., Murphy,B.,
            Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
            Adams,M.D. and Cargill,M.
            Direct submission
            Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
            Rockville, MD 20850, USA
COMMENT   These sequences were made by sequencing genomic exons and ordering
            them based on alignment
FEATURES             Location/Qualifiers
            source          1..3976
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10090"
                        <1..>3976
                        /gene="ATRN"
                        /locus_tag="HGM6607"

ORIGIN
Alignment Scores:
Pred. No.:      0      Length:      3976
Score:          5980.00      Matches:      1089
Percent Similarity: 84.14%      Conservative: 25
Best Local Similarity: 82.25%      Mismatches: 210
Query Match:     75.55%      Indels:      0
DB:              29      Gaps:          0

US-09-787-097-12 (1-1429) x AY418589 (1-3976)
Qy      106 ValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpVal 125
Db      2  GTCAACGGGGCGCGCTGCACCGCGCAGTGGCTGTGCCCCACCGGGCTGGGGT 61
Qy      126 GlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheVal 145
Db      62  GGGAGCAATGCCAGCACTCGGGGGCCGCTTCAGNNNNNNNNNNNNNNNNNNNNNNNNNN 121

```

```

Qy      146 ThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGln 165
Db      122  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNG 181
Qy      166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185
Db      182  CCAAAATGAATAATGAGACTTCGCTTCAACCATTTTGTACAGAAITGAGTGGACCAT 241
Qy      186 LeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaPheSerGlyLeu 205
Db      242  TTATATGTTTATGATGGGACTCAATCTACGACCTCTGATTGCTGCTTTAGNNNNNN 301
Qy      206 IleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyr 225
Db      302  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 361
Qy      226 AlaLeuLeuHisPhePheSerAspAlaAlaIatyrAsnLeuThrGlyPheAsnIleThrTyr 245
Db      362  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 421
Qy      246 SerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSer 265
Db      422  NNTTTTGACATGTGTCGAATAATTGCTCAGGCCGAGAGAGTGTAAAGACGACGTAACAGC 481
Qy      266 SerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIlePro 285
Db      482  AGCAGCGCTGTTGAGTGTGAATGTTCTGAATACTGGAAGGGGAGTCGTGTGACATTCCT 541
Qy      286 HisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArg 305
Db      542  CACTGTACACACAACGTGGCTTTCTCTCACCGAGGCATCTGTAAATCAAGACGATACACAGA 601
Qy      306 GlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsn 325
Db      602  GGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 661
Qy      326 GlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLys 345
Db      662  CAGTCTTTTGGACTCGAGAGAGATATCTCGATTAAAGCTTCCAGAGCCTCTCATAAA 721
Qy      346 AlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAsp 365
Db      722  GCTGTGTGCTCAATGGAATATAATGTGGTGTGTGGCGGATATATGTTCACACCATCAGAT 781
Qy      366 TyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSer 385
Db      782  TACAGCATGGTTTTAGCGTATGACCTGACTTCTAGGGAATGGCTTCCACTAAACCATCT 841
Qy      386 ValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyr 405
Db      842  GTCAACAGATGTGTTGTAAGATATGTTCATCTCTTGGCATTACATAAGNNNNNNNNNN 901
Qy      406 MetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGluLeuArgValPheHis 425
Db      902  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 961
Qy      426 IleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValVal 445
Db      962  NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1021
Qy      446 GlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePhe 465
Db      1022 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1081
Qy      466 GlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsn 485
Db      1082 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1141
Qy      486 ThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSer 505
Db      1142  ACATGGAGTATATTACATACCTCAGGTGCTCTTGTGCAAGGGGGTTATGCCACAGTAGT 1201
Qy      506 ValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAla 525

```

Db 1202 GTTTATGATGACGAGCAAGGCTCTGTACGTTTCATGTTGCTACAGGCTTTCAGCGCC 1261  
Qy 526 AsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetThrThr 545  
Db 1262 AACAAATACCGGCTTGCAATGACCTCTACAGATAGATGGATGATCAGATGGAGCC 1321  
Qy 546 IleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThr 565  
Db 1322 ATTCTTAAGACAGCGGATTTTCGGTTACTTGCATACAGCTGTGATAGTGAAGCC 1381  
Qy 566 MetLeuValPheGlyAsnThrHisAsnAspThrSerMetSerHisClyAlaLysCys 585  
Db 1382 ATGCTGGTGTGGAGGGACACACACATGACCTTCATGAGCCAGGTCGCCAAATGC 1441  
Qy 586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTyrSerValLeuProArg 605  
Db 1442 TTCTCTCAGACTTCATGCTTATGACATTTGNNNNNNNNNNNNNNNNNNNNNNNN 1501  
Qy 606 ProAspLeuHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThr 625  
Db 1502 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 1561  
Qy 626 MetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSer 645  
Db 1562 ATGTATGTGTTGGGGCTTCAACAGCCTCTCTCAGTGTGAGCTTGGTCTTTACCTCG 1621  
Qy 646 GluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArg 665  
Db 1622 GAGCAGTGGCATGCAACCGCAGTGAAGCTGTTGTGGCAGCAGGACCTGGTATCCGG 1681  
Qy 666 CysValTyrAsnThrGlySerSerGlnCysIleSerTyrAlaLeuAlaThrAspGluGln 685  
Db 1682 TGCTGTGGGACACACAGTCGTCTCGATGTACTCTCTGGAGTTGGCACTGAAGAACAA 1741  
Qy 686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705  
Db 1742 GCAGAAAAGCTTAAATCAGAGTGTCTTTCTAAAGAACCTTTGACCATGACAGATGTGAC 1801  
Qy 706 GlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTyrCysAsnAsp 725  
Db 1802 CAGCACACAGATGTTTACAGCTGCAAGCCAAATCAATGATGCTGCCATGTTGCAATGAT 1861  
Qy 726 HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGlu 745  
Db 1862 CACTGTGCTCCTGTGAACACAGCTGCACAGAGGCCAGATCTCCATTGCCAAGTATGAG 1921  
Qy 746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCys 765  
Db 1922 AGTGTCCCAAGGATAACCCCATGTACTGCAATTAAGAAACCCAGCTGCAGGAGCTGT 1981  
Qy 766 AlaLeuAspGlnAsnCysGlnTyrGluProArgAsnGlnGluCysIleAlaLeuProGlu 785  
Db 1982 GCCCTAGACAGAACTGCCAGTGGAGGCCCGGAATCAAGAGTGCATCGCCTGCGGAA 2041  
Qy 786 AsnIleCysGlyIleGlyThrHisLeuValGlyAsnSerCysLeuLysIleThrThrAla 805  
Db 2042 AATATCTGTGGCAATGGCTGGCATTTGGTTGGAACTCGTGTCTGAAATCACTACTGCT 2101  
Qy 806 LysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAsnSer 825  
Db 2102 AAGGAGAATATGACAATGCTAAATGTCTCTAGGAACCAATGCTTTTGGCTTCC 2161  
Qy 826 LeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSer 845  
Db 2162 CTCACATCCAGAGAGGTGGAGTTTGTCTTATAGCAGCTTCGATTATGCAATCATCT 2221  
Qy 846 GlnSerMetSerLysLeuThrLeuThrProTyrValGlyLeuArgLysIleAsnValSer 865  
Db 2222 CAAAGTATGCTCAAGCTCACTGACTCCATGGTGTGTTCTTCGGAAGATCAATGTGTCT 2281  
Qy 866 TyrTyrCysTyrGluAspMetSerProPheThrAsnSerLeuLeuGlnTyrMetProSer 885

Db 2282 TACTGGTGCTGGGAGGATATGTCTCCATTACAAATAGTTTGTGTCAGTGGATGCCATCT 2341  
Qy 886 GluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLys 905  
Db 2342 GAGCCAGTATGATGCTGGCTTCTGTGGATCTTGTGAGAGCTAGTACTCGGGGATTAAAG 2401  
Qy 906 AlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSer 925  
Db 2402 GCTGCAACCTGCATCAACCTCTCAATGGCAGGCTCTGTGAAAGGCGCTGCAAAACCAAGT 2461  
Qy 926 AlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGly 945  
Db 2462 GCCAAGCAGTCCCGACACCATGTGCCCTCGGACAGCGTGTGGAGTGCATAGCAGC 2521  
Qy 946 SerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrVal 965  
Db 2522 AGCTCGAGTGCATGTGGTGCAGTAACATGAAGCAGTGTGTGACTCCAATGCTACCTG 2581  
Qy 966 AlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSerThrCysProProGlu 985  
Db 2582 GCCTCTCTCCCTTTTGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCACCTGAA 2641  
Qy 986 AsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyTyrCys 1005  
Db 2642 AATTGCTCTGGCTACTGCTCTGCAGCCATTGCTGGAGCAGCAGGCTGTGGTGGTGT 2701  
Qy 1006 ThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProVal 1025  
Db 2702 ACTGATCTAGCAATACTGGGAAAGGAAATGATTATGAGGCGAGCTATAAAGACCTGTG 2761  
Qy 1026 LysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSer 1045  
Db 2762 AAGATGCCGTACAGGCTCTGCAGGAAATGTATATCCACAGCCCTCTGAACTCCAGC 2821  
Qy 1046 MetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCysProAlaCysGlnCys 1065  
Db 2822 ATGTGTCTAGGACAGCAGATACAACTGGTCTTTCTATTCTGCTGCTGCTGCTGCTG 2881  
Qy 1066 AsnClyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThr 1085  
Db 2882 AACGGACACAGCAAAATGCATCAACAGAGTATCTGTGAGAAAGTGTGAGGACCTGACCCG 2941  
Qy 1086 GlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyLys 1105  
Db 2942 GGCAAGCACTGCGAGACCTGCATATCTGGCTTCTATGTTGACCCGACTAATGAGGCAAA 3001  
Qy 1106 CysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCys 1125  
Db 3002 TGTGAGCCATGCAAGTGCATGGGCGAGCATCTGTGCAACCAACCAACCCGCAAGTGC 3061  
Qy 1126 PheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArg 1145  
Db 3062 TTCTGTACCAACCAAGGTGTCAAGGGGACGAGTGCAGCATGTGTGAGGTAGAAATCGA 3121  
Qy 1146 TyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe 1165  
Db 3122 TACCAGGAAACCTCTCAAGGAACATGTACTATACCTTCTCATTTGATATCAGTTTC 3181  
Qy 1166 ThrPheSerLeuSerGlnGluAsnAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185  
Db 3182 ACCTTTGTAGCTGCTCCAGGAAGACGACCTGCTACTACAGCCATCACTTTTGGGTACT 3241  
Qy 1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205  
Db 3242 CCTGATGAACAAACAGGATTTGGACATGTTTCATCAATGCCCTCCAAACCTTCAACCTC 3301  
Qy 1206 AsnIleThrTyrAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal 1225  
Db 3302 AACATCACTGGGCCACAGCTTCCAGCCGGAACCCAGACTCGAGAGAGGTCCTGTT 3361  
Qy 1226 ValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPhe 1245  
Db 3362 GTTTCAAAACCAACATCAAGGAATACAAAGATAGTCTTCTTAATGAGAAATTGATTTT 3421

QY	1246	ArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLys	1265
Db	3422	CGCAACATCCAAACATCACTCTTTGTTTATGTCAGTAATTCCTTGGCCCATCAA	3481
QY	1266	IleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThr	1285
Db	3482	ATTACATTGGCTTCTCCAGCACACCAACTTCATGGACCTGGTACAGTTCTTCGTGACT	3541
QY	1286	PhePheSerCysPheLeuSerLeuLeuValAlaValValTrpLysIleLysGln	1305
Db	3542	TTCTTCAGTGTTCCTCTCGCTCTTCGTGGCTGCGAGTGGTCTGGAAGATCAAGCAG	3601
QY	1306	SerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSer	1325
Db	3602	AGCTGTTGGCATCCAGCGGAGAGCAACTTCCTGGAGATGCAACAGATGCCAGC	3661
QY	1326	ArgProPheAlaSerValIleAlaLeuGluThrAspGluGluProProAspLeuIle	1345
Db	3662	CGCCCTTTCCTCTGTAACGTTGCTTGGAAACAGATGAAGAACCTCCTCATCTCATT	3721
QY	1346	GlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys	1365
Db	3722	GGGGGAAGTATAAGACCGTCTTANGCCATTCCTGAGCCCTGCTTGGTAACAA	3781
QY	1366	AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProPro	1385
Db	3782	GCAGCATCTCTCTGTAATTCGTAGGCTCCCTCGAGGACTGGAGGAATCCCTCCTCT	3841
QY	1386	GlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIle	1405
Db	3842	GGTCAGTCAGTCTCCTGCTGGCCAGTCCCTGCTGGTGGCATTTCTCAGCAGATGCATA	3901
QY	1406	ValTyrLysGlyLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnPro	1425
Db	3902	GTGTACAGGAGAGTCAGAGCTGTGAAGAACCGAGACAGACGCGCTGCACAGCCT	3961
QY	1426	GlyThrCysIle	1429
Db	3962	GGAACTGCATT	3973
RESULT 3			
AY418588		3671 bp DNA linear	GSS 17-DEC-2003
LOCUS		Pan troglodytes ATRN gene, VIRTUAL TRANSCRIPT, partial sequence,	
DEFINITION		genomic survey sequence.	
ACCESSION		AY418588	
VERSION		AY418588.1	GI:39774548
KEYWORDS		GSS.	
SOURCE		Pan troglodytes (chimpanzee)	
ORGANISM		Pan troglodytes	
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
PUBMED			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			

source	1. 3671	/organism="Pan troglodytes"	
gene	/mol_type="genomic DNA"		
	/db_xref="taxon:9598"		
	<1..3671		
	/gene="ATRN"		
	/locus_tag="HCM6607"		
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	3671
Score:	5018.00	Matches:	950
Percent Similarity:	77.76%	Conservative:	1
Best Local Similarity:	77.68%	Mismatches:	272
Query Match:	63.40%	Indels:	0
DB:	29	Gaps:	0
US-09-787-097-12 (1-1429) x AY418588 (1-3671)			
QY	166	ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis	185
Db	2	CCAAATAGAAATAATGAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGACCAT	61
QY	186	LeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeu	205
Db	62	TTATATGTTTATGATGGGACTCAATTTATGACCCGTAGTTCGTGCATTTAGCAGTTC	121
QY	206	IleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyr	225
Db	122	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	181
QY	226	AlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyr	245
Db	182	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	241
QY	246	SerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSer	265
Db	242	NNTTTGGTATGTCGCGAATACTGCTCAGGCCGAGGAGGTGAAGATCAGTAATAGC	301
QY	266	SerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIlePro	285
Db	302	AGCGATCTGTTGAATGTGAATGTTCTGAAAACTGGAAGGTGAAGCATGTGCATTCCT	361
QY	286	HisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArg	305
Db	362	CACGTACACACACTGTGCTTTCCTCATCGAGCATCTGCAATTCAGTGTATGTCAGA	421
QY	306	GlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsn	325
Db	422	GGATGCTCCTGCTTCTCAGACTGCGAGNNNNCTGGATGTTTCAGTTCTGTACAGCTAAC	481
QY	326	GlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLys	345
Db	482	CAGTCATTTGGACTCGAGGGAATATTCCTAATTAAGCTCCCGAGCATCTCATAAA	541
QY	346	AlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAsp	365
Db	542	GCTGTGCTCAATGGAACATTTATGTGGGTGTTTGGAGGATATATGTTCAACCACTCAGAT	601
QY	366	TyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAspSer	385
Db	602	TATAACATGTTCTAGCGTATGACCTTGCTCTAGGAGTGGTTCACATTAACCGTTCT	661
QY	386	ValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyr	405
Db	662	GTGAACAAATGTGGTGTGTAGATATGGTCATTCCTTGGCATATACAGNNNNNNNNNN	721
QY	406	MetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGluLeuArgValPheHis	425
Db	722	NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN	781
QY	426	IleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValVal	445







QY 1166 ThrPheSerLeuSerGlnGluAspArgGlyTyrThrAlaIleAsnPheValAlaThr 1185  
 Db 3002 ACCTTTAGTCTATCCAGAGATGCTGCTATTACACAGCTATTAAATTTTGTGCTACT 3061  
 QY 1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205  
 Db 3062 CCTGACGACAAACACAGGATTTGGACATGTTTCATCATGCTCCCAAGAAATTTCAACCTC 3121  
 QY 1206 AsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal 1225  
 Db 3122 AACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGCTGCGAGAGAGATGCTGTT 3181  
 QY 1226 ValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPhe 1245  
 Db 3182 GTTTCAAAACCAACATTAGGAGTACAAAGATAGTTTCTCTAATGAGAGTTTGATTTT 3241  
 QY 1246 ArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLys 1265  
 Db 3242 CGCAACCAACCAATATCATCTTCTTTGTTTATGTCAGTAATTTCACTGGGCCATCAA 3301  
 QY 1266 IleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlaPheValThr 1285  
 Db 3302 ATTCAATGCTCTCTCTCNNAACCAATTTTATGACCTGGTACAGTTCTGCTGACT 3361  
 QY 1286 PhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrpLysIleLysGln 1305  
 Db 3362 TTCTTCAGTTGTTCTCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3421  
 QY 1306 SerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnMetAlaSer 1325  
 Db 3422 AGTTGTGTGGCTCCAGAGCTAGAGAGCAATTTTATGACCTGGTACAGTTCTGCTGACT 3481  
 QY 1326 ArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuIle 1345  
 Db 3482 CGTCCCTTGGCTCCGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3541  
 QY 1346 GlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys 1365  
 Db 3542 GGGGGAGTATAAGACTGTCCCAACCAATGCACTGGAGCCATGTTTGGCAACAA 3601  
 QY 1366 AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProPro 1385  
 Db 3602 GCGCTGTCTCTCTGTTGTGAGGCTCTCTGAGGCTGGTGGTGGTGGTGGTGGTGGT 3661  
 QY 1386 GlyGlnSer 1388  
 Db 3662 GGGCAGTCA 3670

RESULT 4  
 BX440935  
 LOCUS  
 DEFINITION BX440935 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
 CS0DF012YLL19 5-PRIME, mRNA sequence.  
 ACCESSION BX440935  
 VERSION BX440935.1 GI:30781889  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 1201)  
 Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 EVRY cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 Library was constructed by Life Technologies, a division of  
 Invitrogen. This sequence belongs to sequence cluster 10212.f For  
 more information about this cluster, see  
 http://www.genoscope.cns.fr/  
 cgi-bin/cluster.cgi?seq=CS0DF012CF10QPl&cluster=10212.f. Contact :

Feng Liang Email : fliang@lifetech.com URL :  
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600  
 Faraday Avenue Genoscope sequence ID : CS0DF012CF10QPl.  
 Location/Qualifiers  
 source  
 1.1201  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0DF012YLL19"  
 /tissue\_type="FETAL BRAIN"  
 /dev\_stage="fetal"  
 /clone\_lib="Homo sapiens FETAL BRAIN"  
 /notes="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
 was primed with a NotI-oligo(dT) primer. Five prime end  
 enriched, double-strand cDNA was digested with Not I and  
 cloned into the Not I and EcoRV sites of the pCMVSPORT 6  
 vector. Library was not normalized."

## ORIGIN

Alignment Scores:  
 Pred. No.: 7,93e-135 Length: 1201  
 Score: 1746.50 Matches: 323  
 Percent Similarity: 90.28% Conservative: 2  
 Best Local Similarity: 89.72% Mismatches: 20  
 Query Match: 22.07% Indels: 19  
 DB: 13 Gaps: 2

US-09-787-097-12 (1-1429) x BX440935 (1-1201)

QY 780 CysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCys 799  
 Db 58 TGCATTGCCCTGCCGAAAATATCTGTGGCATTTGGTGGCATTTGGTGGAACTCATGT 117  
 QY 800 LeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHis 819  
 Db 118 TTGAAATTTACTACTGCCAAGGAGATTATGACATGCTAAATTTGTTCTGTAGAACAC 177  
 QY 820 AsnAlaLeuLeuAlaSerLeuThrGlnLysValGluPheValLeuLysGlnLeu 839  
 Db 178 AATGCCCTTTTGGCTCTCTTACACCTGATGAAGGTAGAATTTGCTCTTAAAGCAGCTG 237  
 QY 840 ArgIleMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeu 859  
 Db 238 CGAATATGAGTCACTCTCAGACATGTCAGCTCACCTTAACCCATGGGTGGGCTT 297  
 QY 860 ArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeu 879  
 Db 298 CGGAAGATCAATGTGCTACTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTA 357  
 QY 880 LeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluPro 899  
 Db 358 CTACAGTGGATGCGCTGAGCCAGTGTGCTGGGATTTCTGTGGAATTTTATCAGAACCC 417  
 QY 900 SerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGlu 919  
 Db 418 AGTACTCGGGACTGAAGGCTGCACTGCATCAACCCACTCAATGCTGTCTGTAA 477  
 QY 920 ArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939  
 Db 478 AGGCTCTCAACACACAGTGTGAAGCAGTGCAGACACCATGTCCTTGGAGCAGCATGT 537  
 QY 940 GlyAspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959  
 Db 538 CGAATTTGACACAGGCGCAGCTCTGAGTCATGCTGGTGCAGCAACATGAAGCAGTGTG 597  
 QY 960 AspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpThrMet 979  
 Db 598 GACTCCAAATGCTATGTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 657  
 QY 980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999  
 Db 658 AGCACTGCCCCCTGAAATTTGTCAGCTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 717

QY 1000 ProGlyCysGlyTTPCyseThrasProSerAsnThrGlyLysGlyLysCysSileGluGly 1019  
 Db 718 CCAGGCTGGCTGGTGTACTGATCCAGCAATACTGCAAGGGAATGATAGAGGT 777  
 QY 1020 SerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhyTyrProGln 1039  
 Db 778 TCCTATAAGGACCAAGTGAAGTGCCTTCGAAGCCCTACAGGAAATTTCTATCCACAG 837  
 QY 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheileHis 1059  
 Db 838 CCCCTGCTCAATCCAGCATGTGCTAGAGACAGCAGATCAACTGCTCTTTCATTCAC 897  
 QY 1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysSileAsnGlnSerLysCysGluLys 1079  
 Db 898 TGTCAGGCTTGCAATGCAAGCGGCACAGTAAATGCATCAATCAGAGCATCTGTGAGAAG 957  
 QY 1080 CysGluLeuAsnLeuThrGlyLysHisCysGluThrCysSileSerGlyPheTyrGlyAsp 1099  
 Db 958 TGTGAGAACYGACAWA-GMWAG-CACTGCGAGAACTGCATATCTGGTCTC-TACGGKGAT 1014  
 QY 1100 ProThrAsnGlyGlyLysCys-----GlnProCysLysCysAsnGlyHis 1114  
 Db 1015 CCACC-AATGAGGGAATGTMECMASARTGCAWGGGCACCSTCTTTTCAACAACAWACG 1073  
 QY 1115 AlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGly 1134  
 Db 1074 GGAATKCTT-----TKRCACAAAGSGGTWAGGGA 1103

RESULT 5  
 BU506373  
 LOCUS  
 DEFINITION AGENCOURT\_10015485 NIH\_MGC\_94 Mus musculus cDNA clone IMAGE:649441  
 5', mRNA sequence.  
 ACCESSION BU506373  
 VERSION BU506373.1 GI:22812606  
 KEYWORDS EST.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus  
 REFERENCE 1 (bases 1 to 964)  
 AUTHORS NIH-MGC <http://mgi.mcg.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
 Tissue Procurement: The Cepko Laboratory  
 cDNA Library Preparation: Life Technologies, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:  
<http://image.llnl.gov>  
 Plate: LLAM14050 row: k column: 18  
 High quality sequence stop: 592.

FEATURES  
 Location/Qualifiers  
 1..964  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10090"  
 /clone="IMAGE:649441"  
 /tissue\_type="retina"  
 /lab\_host="PH108 (phage-resistant)"  
 /clone\_lib="NIH\_MGC\_94"  
 /note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

ORIGIN  
 Alignment Scores:  
 Pred. No.:

2.47e-124 Length: 964

Score: 1619.00 Matches: 289  
 Percent Similarity: 93.48% Conservative: 12  
 Best Local Similarity: 89.75% Mismatches: 19  
 Query Match: 20.45% Indels: 2  
 DB: 13 Gaps: 0

US-09-787-097-12 (1-1429) x BU506373 (1-964)

QY 666 CysValTyrAsnThrGlySerSerGlnCysSileSerTyrAlaLeuAlaThrAspGluGln 685  
 Db 1 TGCTGTGGGACACACAGTCGTCTCGATGTACCTCTGGGAGTTGGCAATGAAGAACA 60  
 QY 686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705  
 Db 61 GCAGAAAGAGTAAATACAGAGTCTTTTCTAAAGAACCTTGACCATGACAGATGTGAC 120  
 QY 706 GlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTyrCysAsnAsp 725  
 Db 121 CAGCACACAGATTGTACAGCTGCACAGCCATACCATGACTGCCCATGTGTCATGAT 180  
 QY 726 HisCysValProArgAsnHisSerCysSerGluGlyGlnHisSerLysPheArgTyrGlu 745  
 Db 181 CACTGTGTCCTGTGAACACACAGCTGCACAGAGAGGCCAGATCTCCATTGCCAAGTAGAG 240  
 QY 746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysAspSerCys 765  
 Db 241 AGTTGCCCAAGGATACCCCATGTACTCTGCAATAGAAACACAGCTGCAGAGACTGT 300  
 QY 766 AlaLeuAspGlnAsnCysGlnTyrGluProArgAsnGlnGluCysSileAlaLeuProGlu 785  
 Db 301 GCCTAGACACAGAACTGCCAGTGGAGGCCCGGAATCAAGAGTGCATGGCCTGCCGAA 360  
 QY 786 AsnLysCysGlyLeuGlyTyrHisLeuValGlyAsnSerCysLeuLysLysThrAla 805  
 Db 361 AATATCTGTGCAATGGCTGGCAATTTGGTTGGAACCTCGGTCTGAAAATCATCTACTGCT 420  
 QY 806 LysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuAlaSer 825  
 Db 421 AAGGAGAAATATGACATGCTAAATGTCTGTAGGACACCAATGCCCTTTTGGCTTC 480  
 QY 826 LeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgLysMetGlnSerSer 845  
 Db 481 CTCACATCCAGAGAGAGTGGAGTTGTCTTAAGCAGCTTCGATTAATGCAATCATCT 540  
 QY 846 GlnSerMetSerLysLeuThrLeuThrProTyrValGlyLeuArgLysLysLeuValSer 865  
 Db 541 CAAGATATGCCAAGCTCACTCTGACTCCATGGGTGGTCTTCGGAAGATCAATGTGCT 600  
 QY 866 TyrTyrCysTyrGluAspMetSerProPheThrAsnSerLeuLeuGlnTyrMetProSer 885  
 Db 601 TACTGTGTGGGAGGATATGTCTCCATTCAAAATAGTTTGTGTCAGTGCATGCCATCT 660  
 QY 886 GluProSerAspAlaGlyPheCysGlyLysLeuSerGluProSerThrArg-GlyLeuLys 905  
 Db 661 GAGCCCATGATGCTGCTCTGTGGATCTTGTGAGCTTGTGAGCTAGTACTCGGGGATTA 720  
 QY 905 sAlaAlaThrCysSileAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSe 925  
 Db 721 GGCTGCACCTGCATCAACCTCTCAATGCGAGCGTCTGAAAAGSCTCAACACACAG 780  
 QY 925 rAlaLysGlnCysArgThrProCysAlaLeuArgThrAla-CysGlyAspCysThrSerG 945  
 Db 781 TGCCAAGCAGTGGCGGACACCATGTGCTGCCCTGCGGACANCCCTTGTGGCAGTGCATAGCA 840  
 QY 945 LysSerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrV 965  
 Db 841 CCAGCTCGGAGTGCATGTGTGTCAGTACATGAGCAATGTGGGACTCCAAATGCCCTACG 900  
 QY 965 alAlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSerThrCysProG 985  
 Db 901 TGGCTCTCTTCNTTTTNGCCAAANGNATTGAAAGGTATACGATGAGCAGCTGTCCAAC 960  
 QY 985 lu 985

```

Db          961 AA 962

RESULT 6
CB520374    840 bp mRNA linear EST 09-JUL-2003
LOCUS       UI-M-GIO-cel-1-18-0-UI.r1 NIH_BMAP_GIO Mus musculus cDNA clone
DEFINITION  IMAGE:6840067 5', mRNA sequence.
ACCESSION   CB520374
VERSION     CB520374.1 GI:29353729
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 840)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: Dr. Jim Lin, University of Iowa
            cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
            cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
            DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
            Clone Distribution: Distribution information can be found at
            http://genome.uiowa.edu/distribution/mousefl.html
            This clone was contributed by the Brain Molecular Anatomy Project
            (BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..840
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:6840067"
/tissue_type="whole brain"
/dev_stage="embryo 13.5,14.5,16.5,17.5dpc"
/lab_host="DH10B (T1 phage resistant)"
/clone_lib="NIH_BMAP_GIO"
/note="Organ: Brain; Vector: pYX- Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonaldo, Lennon and Soares, Genome Research, 6:791-805,
1996. Denatured RNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with oligo-dT
primer containing a Not I site. Double strand cDNA was
size selected according to mRNA size fraction, ligated
with EcoR I adaptor, digested with NotI and then cloned
directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is ACCGACACAG. This library was created for the University
Iowa Brain Anatomy Project (BMAP). Gene discovery in the
Developing Mouse Nervous System", supported by National
Institute of Mental Health (NIMH), Hemin Chin, Ph.D.,
program coordinator."

ORIGIN
Alignment Scores:
Pred. No.: 9,78e-114 Length: 840
Score: 1491.00 Matches: 262
Percent Similarity: 96.07% Conservative: 7
Best Local Similarity: 93.57% Mismatches: 11
Query Match: 18.84% Indels: 1
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x CB520374 (1-840)

QY 718 AsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGly 737
DB 2 AATGACTGCCACTGGTGCATGATCACTGTGCTGCTGACACACACAGTGCACAGAGGC 61
QY 738 GlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsn 757

```

```

Db          961 AA 962

62 CAGATCTCCATTCGCAAGTATGAGAGTTGCCCAAGGATAACCCCATGTACTACTGCAAT 121
QY 758 LysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsn 777
DB 122 AAGAAAAACCAAGCTGCGAGAGCTGTGCCCTAGACACAGAACTGCCAGTGGAGCCCCGGAAT 181
QY 778 GlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsn 797
DB 182 CAAGAGTGCATCGCCCTGCCGGAATATCTGTGGCAATGCTGGTGGCATTTGGTTGGAAC 241
QY 798 SerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArg 817
DB 242 TCGTGTCTGAAATCACTACTGCTAAGGAGAAATATGACAAATCTAAATTTGCTCTTAG 301
QY 818 AsnHisAsnAlaLeuLeuAlaSerLeuThrGlnLysLysValGluPheValLeuLys 837
DB 302 AACCAATGCTTTTGGCTTCCCTCACATCCAGAGAGGTGGAGTTGCTCTTAG 361
QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
DB 362 CAGCTTCGATTAATGCAATCATCTCAAGATATGCTCAAGCTCACTGCTCATCGGTT 421
QY 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
DB 422 GGCTTCGGAAGATCAATGTCCTTACTGTCTGGAGGATATGCTCAATTCACAAAT 481
QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
DB 482 AGTTTGTGTCAGTGGATGCCATCTGAGCCAGTGTGCTGGCTTCTGGATCTTGTCA 541
QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
DB 542 GAGCCTAGTACTCGGGATTAAAGGCTGCAACCTGATCAACCTCTCAATGCACGCTC 601
QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
DB 602 TGTGAAGGCTGCANACCACAGTGCAGCAGTGGCGACACCATGTGCCCTGCGGACA 661
QY 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
DB 662 GCGTGTGGCGAGTGCATAGCAGCAGCTCGAGTGCATGTGTGTCGAGTACATGAAGCAG 721
QY 958 CysValAspSerAsnAlaTyrValAlaSerPhePheGlyGlnCysMetGluTrpTyr 977
DB 722 TGTGTGACTCCATGCTTACGTTGGCTCTTCCCTTTGGCCAGTGTATGGATGTAT 781
QY 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
DB 782 ACGATGACAGCTGCCACCTGA-NATTGCTCTGGCTACTGTACTACCTGCAGCCATTGCTTG 840

RESULT 7
BG678679    933 bp mRNA linear EST 01-MAY-2001
LOCUS       602624477F1 NCI_CGAP_Skn4 Homo sapiens cDNA clone IMAGE:4749427 5',
DEFINITION  mRNA sequence.
ACCESSION   BG678679
VERSION     BG678679.1 GI:13910076
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 933)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgabs-remail.nih.gov
            Tissue Procurement: James Cleaver, M.D.
            cDNA Library preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
            DNA Sequencing by: Incyte Genomics, Inc.

```

FEATURES	SOURCE
1. <b>Feature 1</b>	Source 1
2. <b>Feature 2</b>	Source 2
3. <b>Feature 3</b>	Source 3
4. <b>Feature 4</b>	Source 4
5. <b>Feature 5</b>	Source 5
6. <b>Feature 6</b>	Source 6
7. <b>Feature 7</b>	Source 7
8. <b>Feature 8</b>	Source 8
9. <b>Feature 9</b>	Source 9
10. <b>Feature 10</b>	Source 10
11. <b>Feature 11</b>	Source 11
12. <b>Feature 12</b>	Source 12
13. <b>Feature 13</b>	Source 13
14. <b>Feature 14</b>	Source 14
15. <b>Feature 15</b>	Source 15
16. <b>Feature 16</b>	Source 16
17. <b>Feature 17</b>	Source 17
18. <b>Feature 18</b>	Source 18
19. <b>Feature 19</b>	Source 19
20. <b>Feature 20</b>	Source 20
21. <b>Feature 21</b>	Source 21
22. <b>Feature 22</b>	Source 22
23. <b>Feature 23</b>	Source 23
24. <b>Feature 24</b>	Source 24
25. <b>Feature 25</b>	Source 25
26. <b>Feature 26</b>	Source 26
27. <b>Feature 27</b>	Source 27
28. <b>Feature 28</b>	Source 28
29. <b>Feature 29</b>	Source 29
30. <b>Feature 30</b>	Source 30
31. <b>Feature 31</b>	Source 31
32. <b>Feature 32</b>	Source 32
33. <b>Feature 33</b>	Source 33
34. <b>Feature 34</b>	Source 34
35. <b>Feature 35</b>	Source 35
36. <b>Feature 36</b>	Source 36
37. <b>Feature 37</b>	Source 37
38. <b>Feature 38</b>	Source 38
39. <b>Feature 39</b>	Source 39
40. <b>Feature 40</b>	Source 40
41. <b>Feature 41</b>	Source 41
42. <b>Feature 42</b>	Source 42
43. <b>Feature 43</b>	Source 43
44. <b>Feature 44</b>	Source 44
45. <b>Feature 45</b>	Source 45
46. <b>Feature 46</b>	Source 46
47. <b>Feature 47</b>	Source 47
48. <b>Feature 48</b>	Source 48
49. <b>Feature 49</b>	Source 49
50. <b>Feature 50</b>	Source 50
51. <b>Feature 51</b>	Source 51
52. <b>Feature 52</b>	Source 52
53. <b>Feature 53</b>	Source 53
54. <b>Feature 54</b>	Source 54
55. <b>Feature 55</b>	Source 55
56. <b>Feature 56</b>	Source 56
57. <b>Feature 57</b>	Source 57
58. <b>Feature 58</b>	Source 58
59. <b>Feature 59</b>	Source 59
60. <b>Feature 60</b>	Source 60
61. <b>Feature 61</b>	Source 61
62. <b>Feature 62</b>	Source 62
63. <b>Feature 63</b>	Source 63
64. <b>Feature 64</b>	Source 64
65. <b>Feature 65</b>	Source 65
66. <b>Feature 66</b>	Source 66
67. <b>Feature 67</b>	Source 67
68. <b>Feature 68</b>	Source 68
69. <b>Feature 69</b>	Source 69
70. <b>Feature 70</b>	Source 70
71. <b>Feature 71</b>	Source 71
72. <b>Feature 72</b>	Source 72
73. <b>Feature 73</b>	Source 73
74. <b>Feature 74</b>	Source 74
75. <b>Feature 75</b>	Source 75
76. <b>Feature 76</b>	Source 76
77. <b>Feature 77</b>	Source 77
78. <b>Feature 78</b>	Source 78
79. <b>Feature 79</b>	Source 79
80. <b>Feature 80</b>	Source 80
81. <b>Feature 81</b>	Source 81
82. <b>Feature 82</b>	Source 82
83. <b>Feature 83</b>	Source 83
84. <b>Feature 84</b>	Source 84
85. <b>Feature 85</b>	Source 85
86. <b>Feature 86</b>	Source 86
87. <b>Feature 87</b>	Source 87
88. <b>Feature 88</b>	Source 88
89. <b>Feature 89</b>	Source 89
90. <b>Feature 90</b>	Source 90
91. <b>Feature 91</b>	Source 91
92. <b>Feature 92</b>	Source 92
93. <b>Feature 93</b>	Source 93
94. <b>Feature 94</b>	Source 94
95. <b>Feature 95</b>	Source 95
96. <b>Feature 96</b>	Source 96
97. <b>Feature 97</b>	Source 97
98. <b>Feature 98</b>	Source 98
99. <b>Feature 99</b>	Source 99
100. <b>Feature 100</b>	Source 100

Alignment Scores:

US-09-787-097-12 (1-1429) x BG678679 (1-933)

QY	1000	ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyysGlyIysCysIleGluGly	1019
DB	3	CCAGGCTGTGGCTGGTGTACTGTATCCAGACAATACTGGCAAGGAAATGCATAGAGGT	62
QY	1020	SerTyrlsGlyProValIysMetProSerGlnAlaProThrGlyAsnPheTyProGln	1039
DB	63	TCCTATAAAGGACCCAGTGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAG	122
QY	1040	ProLeuLeuAsnSerMetCysLeuGluAspSerArgTyAsnTrpSerPheIleHis	1059
DB	123	CCCTGTCTCAATTCACGATGTGTCTAGAGACAGCAGATACAACCTGGTCTTTCAATTCAC	182
QY	1060	CysProAlaCysGlnCysAsnGlyHisSerIysCysIleAsnGlnSerIleCysGluLys	1079
DB	183	TGTCCAGCTTGCCAATGCAACGGGCCACAGTAAATGTCATCAATCAGAGCATCTGTGAGAAG	242
QY	1080	CysGluAsnLeuThrThrGlyIysHisCysGluThrCysIleSerGlyPheTyGlyAsp	1099
DB	243	TGTGAGAACTGACCACAGCAGCAGCATCTGGAGACCTGCATATCTGGCTTCTACGGTGAT	302
QY	1100	ProThrAsnGlyIlyIysCysGlnProCysIysCysAsnGlyHisAlaSerLeuCysAsn	1119
DB	303	CCCACCAATGGAGGGAATGTCAACCATGCAAGTGCATGGGCACGGGTCTCTGTGCAAC	362
QY	1120	ThrAsnThrGlyIysCysPheCysThrThrIysGlyValIysGlyIysAspGluCysGlnLeu	1139
DB	363	ACCAACACGGGCAAGTCTTCTGCACCACCAAGGGCGTCAAGGGGACGAGTGCACGCTA	422
QY	1140	CysGluValGluAsnArgTyGlnGlnIysAsnProLeuHrGlyThrCysTyTyThrIleu	1159
DB	423	TGTGAGGTGAAAAATCGATACCAAGGAAACCCCTCTCAGAGAACATGTATTATACTCTT	482
QY	1160	LeuIleAspTyGlnPheThrPheSerLeuSerGlnGluAspAspArgTyTyThrAla	1179
DB	483	CTTATTGACTATCATGTTTCCACTTTAGTCTATCCCAGGAAGATGATCGCTATTACACAGCT	542
QY	1180	IleAsnPheValAlaThrProAspGluIleAsnArgAspLeuAspMetPheIleAsnAla	1199
DB	543	ATCAATTTTGTGGGTACTCTCTGACGAAACAAACAGGGATTCGGACATGTTTCATCAATGCC	602
QY	1200	SerIysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAla	1219
DB	603	TCAAGAATTTCAACCTCAACATCACTGGGCTCCAGTTTCTCAGCTGGAAACCCAGGCT	662



QY 1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeu---CysGlu 1141  
 Db 1154 AAGTGTGTTTGCACACCAAGGGATCAAGGGTGCACCTGCGAGCTGTATTGTCTAG 1210  
 RESULT 9  
 BY761276  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 REFERENCE  
 AUTHORS  
 Okazaki,Y., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,  
 Nikaido,I., Osato,N., Saito,R., Suzuki,H., Yamanaka,I.,  
 Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,  
 Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,  
 Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,  
 Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,  
 Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,I.,  
 Gariboldi,W., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,  
 Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,  
 Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,  
 Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,  
 Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,  
 Numata,K., Okido,T., Pavan,W.J., Perteau,G., Pesole,G.,  
 Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,  
 Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,  
 Sadelain,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,  
 Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,  
 Verardo,R., Wagner,L., Wahlstedt,C., Wang,Y., Watanabe,Y.,  
 Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,  
 Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,  
 Hayatsu,N., Hirozane-Kishikawa,T., Konno,H., Nakamura,M.,  
 Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,  
 Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,  
 Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,  
 Shinagawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,  
 Rogers,J., Birney,E. and Hayashizaki,Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)  
 1245683  
 1246681  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gsc.riken.go.jp,  
 URL:http://genome.gsc.riken.go.jp/  
 Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,  
 Fukuda,S., Hashizume,W., Hayashida,K., Hirozane,T., Hori,F.,  
 Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kawai,J., Kojima,Y.,  
 Kondo,S., Konno,H., Koya,S., Miyazaki,A., Murata,M., Nakamura,M.,  
 Nomura,K., Numazaki,R., Ohno,M., Ohsato,N., Saito,R., Sakazume,N.,  
 Sano,H., Sasaki,D., Sato,K., Shibata,K., Shiraki,T., Tagami,M.,  
 Takeda,Y., Waki,K., Watahiki,A., Muramatsu,M. and Hayashizaki,Y.  
 Direct Submission  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)  
 Computer-based methods for the mouse full-length cDNA  
 Encyclopedia: real-time sequence clustering for construction of a  
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 cDNA library was prepared and sequenced in Mouse Genome  
 Encyclopedia Project of Genome Exploration Research Group in Riken  
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
 Division of Experimental Animal Research in Riken contributed to  
 prepare mouse tissues.  
 Tissues were provided by David A. Hume (Depts. of Biochemistry  
 and Microbiology/Parasitology Institute for Molecular Bioscience  
 University of Queensland Brisbane, Q 4072 Australia) whose  
 assistance we gratefully acknowledge.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
 further details.

# FEATURES

Location/Qualifiers  
 1..770  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="C57BL/6J"  
 /db\_xref="taxon:10090"  
 /clone="G530008P22"  
 /tissue\_type="bone marrow"  
 /cell\_type="macrophage"  
 /clone\_lib="RIKEN full-length enriched, bone marrow  
 macrophage"

## ORIGIN

Alignment Scores:  
 Pred. No.: 6,07e-106 Length: 770  
 Score: 1397.00 Matches: 240  
 Percent Similarity: 96.09% Conservative: 6  
 Best Local Similarity: 93.75% Mismatches: 10  
 Query Match: 17.65% Indels: 0  
 DB: 13 Gaps: 0

US-09-787-097-12 (1-1429) x BY761276 (1-770)

QY 837 LysGlnLeuArgLysMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTTP 856  
 Db 3 AAGCAGGTGGATGATGCGCATCTCAATGATGATGCGCAGCTCACTGATTCATGG 62  
 QY 857 ValGlyLeuArgLysLeuValSerTyrTyrCysTTPGluAspMetSerProPheThr 876  
 Db 63 GTTGGTCTTCGGAGATCAATGATGCTTACTGGTCTGGGAGGATATGCTCCATTCA 122  
 QY 877 AsnSerLeuLeuGlnTTPMetProSerGluProSerAspAlaGlyPheCysGlyLeu 896  
 Db 123 AATAGTTTGTGTCAGTGGATGTCATCTAGCCAGTGTGCTGCTTCTGGGATCTGN 182  
 QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
 Db 183 TCAGAGCNTATGATCTCGGGATTAAGGCTGCAACCTGCATCAACCTCTCAATGGCAG 242  
 QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
 Db 243 GTCTGTGAAGGCTGCAACACCATGTCACAGTGGCGGACCATGTCCTGCTGGCG 302  
 QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTTPCysSerAsnMetLys 956  
 Db 303 ACAGCGTGTGGCAGTGCATAGCAGCAGCTCGGAGTGCATGTGGTGCAGTAACATCAAG 362  
 QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTTP 976  
 Db 363 CAGTGTGTGACATCCATGCTACGTGGGCTCTCTTCCCTTTTGGCAGTGTATGGATGG 422  
 QY 977 TyrThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
 Db 423 TATACGATGACAGCTGCCACCTGAAAAATGCTCTGGCTACTGTACCTGCAGCCATTGC 482  
 QY 997 LeuGlnProGlyCysGlyTTPCysThrAspProSerAsnThrGlyLysGlyLysCys 1016

483 TTGGAGCAGCCAGGCTGTGGTGTACTGATCTAGCAATATCTGGGAAGGAAATGT 542  
1017 ILEGLUGLYSerTyrIysGlyProValIysMetProSerGlnAlaProThrGlyAsnPhe 1036  
543 ATTGAGGGCAGCTATTAAGAGCCTGTGAAGATGCCGTACAGGCTCTGCAGGAATGTG 602  
1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056  
603 TATCCACAGCCCTTCTGAACCTCCAGCATGTCTGTAGAGCAGCAGATACATCTGGTCT 662  
1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnSerIle 1076  
663 TTCATTCACTGCCAGCTTGCAGTGCACAGCAGCAGCAATGTCATCACCAGAGTATC 722  
1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCys 1092  
723 TGTGAGATGTGAGGACCTGACCAGGCGCAGCACTGGCAGACCTGC 770

BY761390 735 bp mRNA linear EST 17-DEC-2002  
musculus cDNA clone G530012E16 3', mRNA sequence.

BY761390  
BY761390.1 GI:27197002  
EST.  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 735).

OKazaki,I., Furuno,M., Kasukawa,T., Adachi,J., Bono,H., Kondo,S.,  
Nikaido,I., Oseato,N., Saito,R., Suzuki,H., Yamanaka,I.,  
Kiyosawa,H., Yagi,K., Tomaru,Y., Hasegawa,Y., Nogami,A.,  
Schonbach,C., Gojobori,T., Baldarelli,R., Hill,D.P., Bult,C.,  
Hume,D.A., Quackenbush,J., Schriml,L.M., Kanapin,A., Matsuda,H.,  
Batalov,S., Beisel,K.W., Blake,J.A., Bradt,D., Brusic,V.,  
Chothia,C., Corbani,L.E., Cousins,S., Dalla,E., Dragani,T.A.,  
Fletcher,C.F., Forrest,A., Frazer,K.S., Gaasterland,T.,  
Gariboldi,M., Gissi,C., Godzik,A., Gough,J., Grimmond,S.,  
Gustincich,S., Hirokawa,N., Jackson,I.J., Jarvis,E.D., Kanai,A.,  
Kawaji,H., Kawasawa,Y., Kedzierski,R.M., King,B.L., Konagaya,A.,  
Kurochkin,I.V., Lee,Y., Lenhard,B., Lyons,P.A., Maglott,D.R.,  
Maltais,L., Marchionni,L., McKenzie,L., Miki,H., Nagashima,T.,  
Numata,K., Okido,T., Pavan,W.J., Perlea,G., Pesole,G.,  
Petrovsky,N., Pillai,R., Pontius,J.U., Qi,D., Ramachandran,S.,  
Ravasi,T., Reed,J.C., Reed,D.J., Reid,J., Ring,B.Z., Ringwald,M.,  
Sandelin,A., Schneider,C., Semple,C.A., Setou,M., Shimada,K.,  
Sultana,R., Takenaka,Y., Taylor,M.S., Teasdale,R.D., Tomita,M.,  
Verardo,R., Wagner,L., Wahlestedt,C., Wang,Y., Watanabe,Y.,  
Wells,C., Wilming,L.G., Wynshaw-Boris,A., Yanagisawa,M., Yang,I.,  
Yang,L., Yuan,Z., Zavolan,M., Zhu,Y., Zimmer,A., Carninci,P.,  
Hayatsu,N., Hirozane-Kishikawa,T., Kono,H., Nakamura,M.,  
Sakazume,N., Sato,K., Shiraki,T., Waki,K., Kawai,J., Aizawa,K.,  
Arakawa,T., Fukuda,S., Hara,A., Hashizume,W., Imotani,K., Ishii,Y.,  
Itoh,M., Kagawa,I., Miyazaki,A., Sakai,K., Sasaki,D., Shibata,K.,  
Shingawa,A., Yasunishi,A., Yoshino,M., Waterston,R., Lander,E.S.,  
Rogers,J., Birney,E. and Hayashizaki,Y.

Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

22354681  
12466851

Contact: Yoshihide Hayashizaki  
Laboratory for Genome Exploration Research Group, RIKEN Genomic  
Sciences Center(GSC), Yokohama Institute  
The Institute of Physical and Chemical Research (RIKEN)  
1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
Tel: 81-45-503-9222  
Fax: 81-45-503-9216  
Email: genome-res@gsr.riken.go.jp/  
URL: http://genome.gsc.riken.go.jp/  
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Carninci,P.,

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,  
Imotani, K., Ishii, Y., Itoh, M., Kawai, J., Kawai, J., Kojima, Y.,  
Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,  
Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,  
Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,  
Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.  
Direct Submission  
Computational Analysis of Full-Length Mouse cDNAs Compared with  
Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new  
genes. Genome Res. 10 (10), 1617-1630 (2000)  
RIKEN integrated sequence analysis (RISA) system--384-format  
sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
10 (11), 1757-1771 (2000)  
Computer-based methods for the mouse full-length cDNA  
encyclopedia: real-time sequence clustering for construction of a  
nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
cDNA library was prepared and sequenced in Mouse Genome  
Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by David A. Hume ( Depts. of Biochemistry  
and Microbiology/Parasitology Institute for Molecular Bioscience  
University of Queensland Brisbane, Q 4072 Australia ) whose  
assistance we gratefully acknowledge.  
Please visit our web site (<http://genome.gsc.riken.go.jp>) for  
further details.  
Location/Qualifiers  
source  
1. 735  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="G530012E16"  
/tissue\_type="bone marrow"  
/cell\_type="macrophage"  
/clone\_lib="RIKEN full-length enriched, bone marrow  
macrophage"

ORIGIN  
Alignment Scores:  
Pred. No.: 4,76e-102 Length: 735  
Score: 1350.00 Matches: 234  
Percent Similarity: 96.73% Conservative: 3  
Best Local Similarity: 95.51% Mismatches: 8  
Query Match: 17.06% Indels: 1  
DB: 13 Gaps: 0

US-09-787-097-12 (1-1429) x BY761390 (1-735)

QY 848 MetSerLysLeuThrLeuThrProTyrValGlyLeuArgLysIleAsnValSerTyrTrp 867  
Db 2 ATGTCAAGTCTCACTCTGATCTCATGGGT-GGCTTCGGAGATCAATGTCTTACTGG 60  
QY 868 CysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluPro 887  
Db 61 TGTCTGNAGGATATGCTCCATTACAAATAGTNTGCTGCAGTGGATGCCATCTGAGCCC 120  
QY 888 SerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAla 907  
Db 121 AGTATGATGCTTCTCTGTGGATCTTGTTCAGACCTAGTACTCGGGGATTAAGGCTGCA 180  
QY 908 ThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLys 927  
Db 181 ACCTGTCATCAACCTCTCAATGGCAGCGCTCTGTGAAGGCTCTCAACACCACTGCCAAG 240  
QY 928 GlnCysArgThrProCysAlaLeuAcrThrAlaCysGlyAspCysThrSerGlySerSer 947  
Db 241 CAGTGCCTGACACCATCTGCTCCCTCGGACAGCGGTGTGGCAGTGGCTACTACGACGCTCG 300  
QY 948 GluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSer 967



```

Db      301 GAGTGCATGGTGCAGTACATGAAGCAGTGTGTGACTCCAAATCCCTACGTGGCTCC 360
Qy      968 PheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProGluLeuEnCys 987
Db      361 TTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCACCTGAAATTCG 420
Qy      988 SerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAsp 1007
Db      421 TCGGGTACTGTACCTGCAGCCATCTCTTGGAGCAGCCAGGCTGTGGTGGTACTGAT 480
Qy      1008 ProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMet 1027
Db      481 CCTAGCAATACTGGGAAGGAAATGATATGAGGCGAGCTATTAAGGACCTGTGAGATG 540
Qy      1028 ProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCys 1047
Db      541 CCCTCAGAGCCCTCTCTCGAGAAATGTATATCCACAGCCCTCTCTGAACCTCCAGCATGTGT 600
Qy      1048 LeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGly 1067
Db      601 CTAGAGCAGCAGAGATCAACTGGTCTTTCATCTGCTCCAGCTGCCAGTGCACACGGA 660
Qy      1068 HisSerLysCysIleAsnGlnSerIleCysGlyLysCysGluAsnLeuThrThrGlyLys 1087
Db      661 CACAGCAATGCATCAACAGAGTATCTGTGAGAAGTGTGAGGACCTGACACCGGCGAAG 720
Qy      1088 HisCysGluThrCys 1092
Db      721 CACTGCGAGACCTGC 735

CF745736      777 bp      mRNA      linear      EST 10-OCT-2003
UI-M-GVO-clw-d-20-0-UI.r1 NIH_BMAP_GVO Mus musculus cDNA clone
IMAGE:30620155 5', mRNA sequence.
CF745736
EST.
CF745736.1 GI:37642076
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 777)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgaabs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Distribution information can be found at
http://genome.uiowa.edu/distribution/mousefl.html
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)
Seq primer: pYX-5.
Location/Qualifiers
1..777
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:30620155"
/tissue_type="whole brain"
/dev_stage="1,5, and 15 days newborn"
/lab_host="PH10B (T1 phage resistant)"
/clone_lib="NIH BMAP_GVO"
/note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I;
Site 2: Not I; The library was constructed according
Bonafide, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured RNA was size fractionated on a 1% agarose

```

## FEATURES

source

gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is GGAATGAA. This library was created for the University of Iowa Brain Anatomy Project (BMAP): "Gene Discovery in the Developing Mouse Nervous System", supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

## ORIGIN

```

Alignment Scores:
Pred. No.:      4,41e-99      Length:      777
Score:          1315.00      Matches:      229
Percent Similarity: 93.80%      Conservative: 13
Best Local Similarity: 88.76%      Mismatches:  16
Query Match:      16.61%      Indels:      0
DB:              14          Gaps:        0

US-09-787-097-12 (1-1429) x CF745736 (1-777)

Qy      566 MetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCys 585
Db      3 ATGCTGTGTGTTGAGGGAGACACACATGACACTTCCATGAGCCAGCGTGCCTAATGC 62
Qy      586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArg 605
Db      63 TTCCTCTCAGACTTTCATGGCTTATGACATTGCTGTGACCGATGTCAGTCTTCCAG 122
Qy      606 ProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThr 625
Db      123 CCTGAGTCCCATCATGATGTCACAGATTTGGCCATTTCAGCAGTCTTGTACACAGACCC 182
Qy      626 MetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSer 645
Db      183 ATGTATGTGTTCGGCGGCTTCAACAGACCTCTCTCAGTCAGCGTCTTGGTCTTTACCTCG 242
Qy      646 GluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArg 665
Db      243 GAGCAGTGCAGTGCACACCCAGTGAAGCTGCTGTGTGGCAGCAGGACCTGGTATCCGG 302
Qy      666 CysValTrpAsnThrGlySerSerGlnCysLeuSerTrpAlaLeuAlaThrAspGluGln 685
Db      303 TGTCTGTGGACACACAGTGTCTCGATGATACCTCTCTGGGAGTGGCACTGAGAACAA 362
Qy      686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705
Db      363 GCAGAAAAGTTAAATCAGAGTGTCTTCTTAAAGAACCTTGACCATGACAGATGTGAC 422
Qy      706 GlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAsp 725
Db      423 CAGCACACAGATTGTTACAGTGCACAGCCAAATACCAATGACTGCCACTGGTGGTCAATGAT 482
Qy      726 HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGlu 745
Db      483 CACTGTGTCCCTGTGAACACACAGCTGCACAGAGGCCAGATCTCCATGTCAGATGAG 542
Qy      746 AsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCys 765
Db      543 AGTTGCCCAAGGATAACCCCATGTACTACTGCAATAAGAAAAACACAGCTGCAGAGCTGT 602
Qy      766 AlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnCysIleAlaLeuProGlu 785
Db      503 GCCCTAGACCAAGAACTGCCAGTGGAGCCCCGGGAATCAAGAGTGCATGCCCTTCCCGGAA 662
Qy      786 AsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrAla 805
Db      663 AATATCTGTGCAATGCTGGCATTTTNGTGTGAACTCGTGTCTGAAAAATCACTACTGCT 722
Qy      806 LysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823

```

Db	723	AAGGGAGATTATGACAATGCTAAATTGCTCTGTANNANNCACATGCCTTTTG	776
RESULT 12			
BM783739			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-09-787-097-12 (1-1429) x BM783739 (1-681)			
Qy	838	GlnLeuArgIleMetGlnSerSergInSerMetSerIlysLeuThrProTrpVal	857
Db	1	CAGTCGCAATATGCAATCTCAGAGCATGTCACCTCAACCTCAACCCATGGGTC	60
Qy	858	GlyLeuArgIlysLeuAanValSerTyTrpCysTrpGluAspMetSerProPheThrAan	877
Db	61	GGCCTTCGGAGATCAATGCTCTACTGTGCTGGAGATATGTCCTTACAAAT	120
Qy	878	SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer	897
Db	121	AGTTTACTACAGTGGATGCGCTCTGAGCCCGAGTGTGGATCTGTGGAAATTTATCA	180
Qy	898	GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal	917
Db	181	GAACCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGC	240
Qy	918	CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr	937
Db	241	TGTGAAGGCTCGCAACACACAGTGTAGCAGTGGCGGACACCATGTGCTTGAGGACA	300
Qy	938	AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln	957
Db	301	GCATGTGGAGATTGCACAGCGCAGCTCTGAGTGCATGTGTGGCGAGCAACATGAAGCAG	360
Qy	958	CysValAspSerAsnAlaTyValAlaSerPheProPheGlyGlnCysMetGluTrpTyTr	977
Db	361	TGTGTGAGCTCCAAATGCCTATGTGGCTCTCTCTTTGGCCAGTGATGGATGGTAT	420
Qy	978	ThrMetSerThrCysProProGluAanCysSerGlyTyTrpCysThrCysSerHisCysLeu	997
Db	421	ACGATGAGCACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCTTCTT	480
Qy	998	GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle	1017
Db	481	GAGCAACCGGCTGTGGCTGGTGTACTGATCCAGCAATACTGGCAAGGAAATGCATA	540
Qy	1018	GluGlySerTyTrpLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTy	1037
Db	541	GAGGGTTCTTATAAGGACCGAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTAT	600
Qy	1038	ProGlnProLeuLeuAanSerSerMetCysLeuGluAspSerArgTyTrpAsnTrpSerPhe	1057
Db	601	CCACAGCCCTGTCTCAATTCAGCATGTGTAGAGGACGACGATACAACTGGTCTTTC	660
Qy	1058	IleHisCysProAlaCysGln	1064
Db	661	ATTCATGTCCAGCTTCCCAA	681
RESULT 13			
BM783258			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
FEATURES			
source			
ORIGIN			
Alignment Scores:			
Pred. No.:			
Score:			
Percent Similarity:			
Best Local Similarity:			
Query Match:			
DB:			
US-09-787-097-12 (1-1429) x BM783739 (1-681)			
Qy	838	GlnLeuArgIleMetGlnSerSergInSerMetSerIlysLeuThrProTrpVal	857
Db	1	CAGTCGCAATATGCAATCTCAGAGCATGTCACCTCAACCTCAACCCATGGGTC	60
Qy	858	GlyLeuArgIlysLeuAanValSerTyTrpCysTrpGluAspMetSerProPheThrAan	877
Db	61	GGCCTTCGGAGATCAATGCTCTACTGTGCTGGAGATATGTCCTTACAAAT	120

```
/clone="S5SNU484-31-H08"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10P"
/clone_lib="S5SNU484"
/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI;
Site 2: NotI; The poly (A) + RNA was decapped with tobacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-tailed
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10P' by electroporation
method."
```

601 CCACAGCCCTGCTCAATTCCAGCATGTGCTAGAGACAGACAGATACACTGCTCTTC 660

1058 IleHisCysProAlaCys 1063

661 ATTCACGTGCTCCAGCTTGC 678

RESULT 14

BM783773

LOCUS

DEFINITION K-EST0061775 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-D02 5', mRNA sequence.

ACCESSION

VERSION BM783773.1 GI:19132005

KEYWORDS

SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE

AUTHORS Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R., Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, J.M., Park, H.S., Kim, S. and Kim, Y.S.

TITLE 21C Frontier Korean EST Project 2001

JOURNAL Unpublished (2002)

COMMENT Contact: Kim YS

Genome Research Center

Korea Research Institute of Bioscience & Biotechnology

52 Eoeun-dong Yuseong-gu, Daejeon 305-333, South Korea

Tel: +82-42-860-4470

Fax: +82-42-860-4409

Email: yongsun@mail.kribb.re.kr

Plate: 36 row: D column: 02

High quality sequence stop: 680.

Location/Qualifiers

1. 680

source

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="S5SNU484-36-D02"

/sex="M"

/tissue\_type="Stomach"

/cell\_type="Epithelial"

/cell\_line="SNU-484"

/lab\_host="Top10P"

/lab\_lib="S5SNU484"

/notes="Organ: Stomach; Vector: pTZ18RP1; Site 1: EcoRI; Site 2: NotI; The poly (A) + RNA was decapped with tobacco acid pyrophosphatase (TAP) and ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase. The first strand cDNA was synthesized from oligo dt-tailed mRNA by priming with dt-tailed vector. The dt-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10P' by electroporation method."

Alignment Scores:

Pred. No.: 2,41e-98 Length: 680

Score: 1305.00 Matches: 226

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 16.49% Indels: 0

DB: 12 Gaps: 0

US-09-787-097-12 (1-1429) x BM783258 (1-680)

QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857

DB 1 CAGCTGCCGAATATCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGC 60

QY 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTyrGluAspMetSerProPheThrAsn 877

DB 61 GGCCTTCGGAGATCAATGTGCTCTACTGTGCTGGAGATATGTCCTCCCATTTACAAT 120

QY 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyLeuSer 897

DB 121 AGTTACTACAGTGGATGCGCTGAGCCAGTGTGCTGGATTCTGTGGAATTTATCA 180

QY 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917

DB 181 GAACCCAGTACTCGGGGCTGAAGGCTGCACTGCATCAACCCACTCAATGGTAGTGTGC 240

QY 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937

DB 241 TGTGAAGGCTGCAACACACAGTGTGAAGTGTGCGGACACCATGTGCTTGAGGACA 300

QY 938 AlaCysGlyAspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGln 957

DB 301 GCATGTGGAGATGACACAGCGGAGCTCTGTAGTGCATGTGTGTCAGCAACATGAAGCAG 360

QY 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977

DB 361 TGTGTGGACTCCAGTGCCTATGTGGCTTCTTCCCTTTTGCCAGTGTATGGATGGTAT 420

QY 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997

DB 421 ACCGATGAGCACCTGCCCCCTGAAATTTGTTCCAGGCTACTGTACCTGTAGTCATTGCTTG 480

QY 998 GluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyCysCysIle 1017

DB 481 GAGCAACACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540

QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037

DB 541 GAGGGTTCTATAAGGACCACTGAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 600

QY 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057

QY 838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857

```

Db      1  CAGCTGCGAATATGCAATCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC 60
QY      858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
Db      61  GGCTTTGGAAGATCAATGTCTCTACTGCTGCTGGAGAGATATGCCCCATTACAAT 120
QY      878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
Db      121 AGTTTACTACAGTGGATGCGCTCTGAGCCAGCAGTGTCTGGATCTCTGTGGAAATTTATCA 180
QY      898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
Db      181 GAACCCAGTACTCGGGACTGAAGGCTGCAACTGTGATCAACCCACTCATGCTAGTGTGTC 240
QY      918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
Db      241 TGTGAAGGCTGCAAAACACACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGAAGACA 300
QY      938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
Db      301 GCATGTGGAGATTGCCACCGCGCAGCTCTGAGTGCATGTGTGCAGCAACATGAAGCAG 360
QY      958 CysValAspSerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977
Db      361 TGTGTGGACTCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 420
QY      978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
Db      421 ACATGAGCAGCTGCCCCCTGAAATTTGTTAGGCTACTGTACCTGTACTGCTTGTTC 480
QY      998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
Db      481 GAGCAACACGAGCTGTGGCTGTGTACTGATCCAGCAATCTGGCCAGGAAATGTCATA 540
QY      1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
Db      541 GAGGGTTCCTATTAAGGACAGTGAAGATGCTTCGAAAGCCCTACAGGAATTTCTAT 600
QY      1038 ProGlnProLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
Db      601 CCACAGCCCTGTCTCAATTCAGCATGTGTGTAGAGGACAGCAGATACAACTGGTCTTTC 660
QY      1058 IleHisCysProAlaCys 1063
Db      661 ATTCACTGTCCAGCTTGC 678

RESULT 15
BM783788
LOCUS      K-EST0061794 S5SNU484 Homo sapiens cDNA clone S5SNU484-36-E09 5',
DEFINITION mRNA sequence.
ACCESSION BM783788.1 GI:19132020
VERSION    1
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Homo sapiens
REFERENCE  1 (bases 1 to 683)
AUTHORS   Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          Kim, N.S., Hahn, Y., Oh, J.H., Lee, J.Y., Ahn, H.Y., Chu, M.Y., Kim, M.R.,
          Oh, K.J., Cheong, J.E., Sohn, H.Y., Kim, O.M., Park, H.S., Kim, S. and
          Kim, Y.S.
          21C Frontier Korean EST Project 2001
          Unpublished (2002)
          Contact: Kim YS
          Genome Research Center
          Korea Research Institute of Bioscience & Biotechnology
          52 Eoeun-dong, Yuseong-gu, Daejeon 305-333, South Korea
          Tel: +82-42-860-4470
          Fax: +82-42-860-4409
          Email: yongsung@mail.kribb.re.kr
          Plate: 36 row: E column: 09
          High quality sequence stop: 683.

```

```

FEATURES             source
Location/Qualifiers
1..683
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="S5SNU484-36-E09"
/sex="M"
/tissue_type="Stomach"
/cell_type="Epithelial"
/cell_line="SNU-484"
/lab_host="Top10F"
/clone_lib="S5SNU484"
/note="Organ: Stomach; Vector: pTZ189PL; Site 1: EcoRI;
Site 2: NotI; The poly (A)+ RNA was decapped with tabacco
acid pyrophosphatase (TAP) and ligated with DNA-RNA linker
including EcoRI site by treatment of T4 RNA ligase. The
first strand cDNA was synthesized from oligo dt-selected
mRNA by priming with dt-tailed vector. The dt-tailed
vector was adjusted to have about 60nt. The cDNA vector
was circularized with E. coli DNA ligase after digestion
of EcoRI which site is also included in vector. An RNA
strand converted to a DNA strand by Okayama-Berg method.
The obtained cDNA vectors were used for transformation of
competent cells E. coli Top10F by electroporation
method."
ORIGIN
Alignment Scores:      Length:      683
Pred. No.:            1301.00      Matches:      225
Score:                99.56%      Conservative:  1
Percent Similarity:   99.12%      Mismatches:   1
Best Local Similarity: 16.44%      Indels:       0
Query Match:         12           Gaps:        0
DB:
US-09-787-097-12 (1-1429) x BM783788 (1-683)
QY      838 GlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857
Db      1  CAGCTGCGAATATGCAATCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTC 60
QY      858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877
Db      61  GGCTTTGGAAGATCAATGTCTCTACTGCTGCTGGAGAGATATGCCCCATTACAAT 120
QY      878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897
Db      121 AGTTTACTACAGTGGATGCGCTCTGAGCCAGTGTCTGGATCTCTGTGGAAATTTATCA 180
QY      898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917
Db      181 GAACCCAGTACTCGGGACTGAAGGCTGCAACCTGTCAACCCACTCAATGCTAGTGTGTC 240
QY      918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937
Db      241 TGTGAAGGCTGCAAAACACACAGTGTAAAGCAGTGGCGGACACCATGTGCTTGAAGACA 300
QY      938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957
Db      301 GCATGTGGAGATTGCCACCGCGCAGCTCTGAGTGCATGTGTGCAGCAACATGAAGCAG 360
QY      958 CysValAspSerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977
Db      361 TGTGTGGACTCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAATGGTAT 420
QY      978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997
Db      421 ACATGAGCAGCTGCCCCCTGAAATTTGTTAGGCTACTGTACCTGTACTGCTTGTTC 480
QY      998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017
Db      481 GAGCAACACGAGCTGTGGCTGTGTACTGATCCAGCAATCTGGCCAGGAAATGTCATA 540

```

```

QY 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
541 GAGGGTTCCATAAAGGACCACTGAAGATGCCCTCCCAAGCCCTACAGGAAATTTCTAT 600
QY 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
601 CCACAGCCCTGCTCAATCCAGCAATGTCTAGAGGACAGCAGATACAACTGGTCTTTC 660
QY 1058 IleHisCysProAlaCysGln 1064
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
661 ATTCAGTGTCCAGCTTGCAA 681

```

Search completed: March 1, 2004, 07:06:40  
Job time : 6974 secs

***This Page Blank (uspto)***

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 29, 2004, 11:51:51 ; Search time 1076 Seconds  
(without alignments)  
5641.891 Million cell updates/sec

Title: US-09-787-097-12

Perfect score: 7915

Sequence: 1 MYAAAAAEARLRRTAATA.....KSGAVNRKQPPAQTGTCI 1429

Scoring table:

BLOSUM62	Xgapop 10.0	Xgapext 0.5
Ygapop 10.0	Ygapext 0.5	
Fgapop 6.0	Fgapext 7.0	
Delop 6.0	Delext 7.0	

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame\_p2n.model -DEV=xlp  
-Q=/cgn2\_1/usfto.spool\_p/us09787097/runat\_26022004\_062934\_23410/app\_query.fasta\_1.1607  
-DB=N Geneseq 29Jan04 -OPT=fastp -SUFFIX=ring -MINMATCH=0.1 -DOOPL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blotsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09787097 @CGN 1.1137 @runat\_26022004\_062934\_23410 -NCFU=6 -ICPU=3  
-NO WRAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV\_TIMEOUT=120 -WARN\_TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 29Jan04:\*

1:	Geneseqn1980s:*
2:	Geneseqn1990s:*
3:	Geneseqn2000s:*
4:	Geneseqn2001as:*
5:	Geneseqn2001bs:*
6:	Geneseqn2002s:*
7:	Geneseqn2003as:*
8:	Geneseqn2003bs:*
9:	Geneseqn2003cs:*
10:	Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	7915	100.0	4290	3	Az52276 Human mem
2	7915	100.0	8561	5	Aas72659 DNA encod
3	7696.5	97.2	8390	3	Aaz91920 Human mah
4	7478	94.5	4068	3	Az52275 Human mem
5	7342.5	92.8	8739	9	Adb53249 Primary r
6	7294	92.2	8827	3	Aaz91916 Wild type
7	7099	89.6	3819	3	Aaz52277 Human sol
8	6870.5	86.8	4072	3	Aaz91921 Human mah

9	6612	83.5	3597	3	Az52274	Human sol
10	6612	83.5	3597	5	Aas72657	DNA encod
11	6612	83.5	3597	9	ADD70997	Human att
12	5622	71.0	3490	5	Aas72660	DNA encod
13	4637.5	58.6	4140	7	ABZ58709	Human hom
14	4623	58.4	4134	7	ABZ58710	Murine ho
15	4136.5	52.3	2625	3	Aaz91922	Human mah
16	3880	49.0	3609	6	AB563434	Human CDN
17	3452	43.6	6373	3	Aaz91919	Human mah
18	3435	43.4	2419	3	Aaz91917	Murine ma
19	2506.5	31.7	6517	4	AAH98748	Human EST
20	2413.5	30.5	6532	5	AAS86766	DNA encod
21	2232.5	28.2	4164	4	ABL09747	Drosophil
22	2084	26.3	3165	5	AAS78781	DNA encod
23	2084	26.3	3313	5	AAS76294	DNA encod
24	2032.5	25.7	10315	4	ABL09746	Drosophil
25	1825	23.1	1012	2	AAH84568	Biorhyth
26	1298	16.4	1654	7	ABX34730	Human mdd
27	1259.5	15.9	1383	6	AAL54064	cDNA of h
28	1236.5	15.5	1051	3	Aaz91918	Murine ma
29	977.5	12.3	8875	9	ADC51661	DNA encod
30	977.5	12.3	9921	9	ADC51659	DNA encod
31	950	12.0	6056	6	AB551293	Human REM
32	937	11.8	8865	4	ABL12959	Drosophil
33	923	11.7	7244	7	ABX70916	Novel hum
34	895	11.3	11502	4	ABL12958	Drosophil
35	840	10.6	207433	5	ABZ72040	Gene 316
36	840	10.6	207433	7	ABX74891	BAC1098L2
37	766	9.7	90050	3	Aaz91925	Wild type
38	722	9.1	671	5	AAS72658	DNA encod
39	708	8.9	982	6	ABL39784	Human NS
40	684.5	8.6	657	7	ABZ58714	Upregulat
41	557	7.0	668	4	AAS02409	Human sec
42	553	7.0	292	3	AAA45505	Human sec
43	536	6.8	481	4	AAL44686	Human bre
44	523	6.6	482	4	AAL23548	Human bre
45	477	6.0	845	6	ABQ52394	Oligonuci

#### ALIGNMENTS

RESULT 1

Az52276  
ID Az52276 standard; cDNA; 4290 BP.

XX AC Az52276;

XX DT 18-JUL-2000 (first entry)

XX DE Human membrane attractin-2 cDNA.

XX KW Human; membrane attractin-2; immune response; macrophage; monocyte;  
T cell; immunostimulant; immunosuppressed patient; cancer;  
immunodeficiency syndrome; transplant; autoimmune disease; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX CDS 1..4290

XX FT /\*tag=

XX FT /product= "Membrane attractin-2"

XX PN WO200015651-A1.

XX PD 23-MAR-2000.

XX PF 14-SEP-1999; 99WO-US020948.

XX PR 14-SEP-1998; 98US-O100137P.

XX PA (DAND ) DANA FARBER CANCER INST INC.

XX PI Duke-Cohan JS, Schlossman SP;





QY 541 ThrGlnMetTrpThrIleLeuIysAspSerArgPheArgTyrLeuHisThrAlaVal 560  
Db 1621 ACCGAGATGGACCAATCTTAAGACACAGCCGATTTTCGGTTACTTGCACACAGCTGTG 1680  
QY 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
Db 1681 ATAGTGAGTGGAAACCAATGCTGTGTGTGGGGAAACACACAATGACACATCTATGAGC 1740  
QY 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTyr 600  
Db 1741 CATGGCGCCAAATGCTTCTCTTCAGATTTTCATGGCCTATGACATTTGCGCTGTGACCGCTGG 1800  
QY 601 SerValLeuProArgProAspLeuHisHisaspValAsnArgPheGlyHisSerAlaVal 620  
Db 1801 TCAGTGTCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCACAGATC 1860  
QY 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640  
Db 1861 TTACACAAACAGCACCATGATGTGTTCGGTGGTTTCATAGTCTCTCTCTCAGCAGCATC 1920  
QY 641 LeuValPheThrSerGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660  
Db 1921 CTGGTATTCACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCGCTGTGTTTAGAGCA 1980  
QY 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680  
Db 1981 GGACCTGGTATTCGGTGTGTGGAAACACAGGTTGTCTCAGTGTATCTCGTGGCGCTG 2040  
QY 681 AlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700  
Db 2041 GCAACTGATGAACAAGACAAAGTTAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100  
QY 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720  
Db 2101 CATGACAGATGGACAGACACACAGATGTTACAGCTGTACAGCCAAACACCAATCACTGC 2160  
QY 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
Db 2161 CACTGGTGAATGACCAATGTTGTCCCCAGGAACACACAGCTGTCAGAGGCCAGATCTCC 2220  
QY 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysIysThr 760  
Db 2221 ATTTTATAGTATGAGAAATGCCCCAGGATAACCCCATGTACTACTTAACAAGAAGACC 2280  
QY 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780  
Db 2281 AGCTGACAGAGCTGTGCCCTGACACAGAACTGCCAGTGGAGCCCGGGAATCAGGAGTGC 2340  
QY 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTyrPheHisLeuValGlyAsnSerCysLeu 800  
Db 2341 ATTGCCCTGCCGAAAATATCTGTGGCATTGGCTGGCAATTTGGTTGGAAACTCATGTTTG 2400  
QY 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820  
Db 2401 AAAATTACTACTGCCAAGAGAAATATGACAAATGTCTAAATTTGTTCTGTAGGAACACCAAT 2460  
QY 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840  
Db 2461 GCCCTTTGGCTTCTCTTACACCCAGAGAAGAGTAGAATTTGTCTTAAGCAGCTGCGA 2520  
QY 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpIleValGlyLeuArg 860  
Db 2521 ATAATGCACTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGGCGCTTCGG 2580  
QY 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880  
Db 2581 AAGATCAATGTCTCTACTGTGTGGAGAGATATGTCCCAATTTTACAAATAGTTACTA 2640  
QY 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
Db 2641 CAGTGGATCCGCTCTGAGCCAGTGTGCTGATGCTGGATTCGTGGAAATTTTATCAGAACCCAGT 2700

QY 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
Db 2701 ACTCGGGACTGAAGGCTGCAAACCTGCATCAACCACTCAATGCTAGTGTCTGTGAAAGG 2760  
QY 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
Db 2761 CCTCAAAACACAGTGTAAAGCAGTCCGGACACCATGTGCTTGGAGACAGCATGTGGA 2820  
QY 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960  
Db 2821 GATTGCACACAGCGCAGCTCTGAGTGCATGTGGTGACGAAACATGAAGCAGTGTGTGGAC 2880  
QY 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980  
Db 2881 TCCAAATCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATAGCATGAGC 2940  
QY 981 ThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000  
Db 2941 ACCTGCCCTGAAAATTTGTTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACCA 3000  
QY 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020  
Db 3001 GGCTGTGGCTGGTGTACTGATCCAGCAATACTGGCAAGGAAATGCATAGAGGGTTC 3060  
QY 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040  
Db 3061 TATAAAGACACAGTGAAGATGCCCTTCGCAAGCCCTACAGAAATTTCTATCCACAGCCC 3120  
QY 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060  
Db 3121 CTGCTCAATTCAGCATGTGTAGAGACAGCAGATACAACTGGTCTTTCATTCACCTGT 3180  
QY 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080  
Db 3181 CCAGCTTGCCAAATGCAACGGCCACAGTAAATGTCATCAATCAGAGCATCTGTGAGAAGTGT 3240  
QY 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100  
Db 3241 GAGAACCTGACACACAGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGTATCCC 3300  
QY 1101 ThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120  
Db 3301 ACCAATGAGGGAATATCTCAGCCATGCAAGTGAAGTGGGCACGCTCTCTGTGCAACACC 3360  
QY 1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140  
Db 3361 AACACGGGCAAGTCTTCTGCACCAACAGGGCGTCAAGGGGACGAGTGCAGCTATGT 3420  
QY 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160  
Db 3421 GAGTAGAAAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTCTT 3480  
QY 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrThrAlaIle 1180  
Db 3481 ATTGACTATCAGTTTCACTTTAGTCTATCCAGAAAGATGATCGCTANTTACACAGCTATC 3540  
QY 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200  
Db 3541 AATTTTGTGGCTACTCTCAGCAACAAAACAGGATTTGGACATGTTTCATCAATGCCTCC 3600  
QY 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220  
Db 3601 AAGAAATTCACCTCAACATCCTGGCTGGCAGTTTCTCAGCTGGAAACCCAGGCTGGA 3660  
QY 1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240  
Db 3661 GAAGAGATGCTGTGTTTCAAAAACCAACATTAAGAGGTACAAAGATAGTTTCTCTAAT 3720  
QY 1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260  
Db 3721 GAGAGGTTTGAATTCGCAACACCCCAAAATATCACTTCTTTGTTATGTGAGTAATTC 3780  
QY 1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280



QY 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220  
DB 601 GCATTTAGTGGCTCATTTCTCTGAGAGATGGCAATGAGACTGCTCCCTGAGGTGGT 660  
QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPheSerAspAlaAlaTyrAsnLeuThrGly 240  
DB 661 GCCACATCAGGTATGCTGCTGCATTTTCTAGTGATGCTGCTTATATTTGACTGGA 720  
QY 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260  
DB 721 TTTPAATATTACTTTACAGTTTGTATATGTCTCAATAAATGCTCAGCGCCGAGAGAGTGT 780  
QY 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280  
DB 781 AAGATCAGTAATAGCAGCGAATGTTGAATGTGAATGTTCTGAAAACCTGGAAGGTGNA 840  
QY 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300  
DB 841 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTTTTCTCATCGAGGCATCTGCAAT 900  
QY 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGluGlyProGlyCysSerVal 320  
DB 901 TCAGTGTATGTACAGGATGCTCTCTCTCAGACTGCGAGGTCCTGGATGTTTCAGTT 960  
QY 321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340  
DB 961 CCTGTACCACTAACCACTGTCATTTGGACTCGAGAGGAATATCTAACTTAAAGCTCCCC 1020  
QY 341 ArgAlaSerHisLysAlaValAlaAsnGlyAsnIleMetTrpValValGlyTyrMet 360  
DB 1021 AGAGCATCTCATAAAGCTGTGTCTCAATGGAACATTTATGTGGTGTGTGGAGGATATATG 1080  
QY 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380  
DB 1081 TTCAACCACTCAGATTATAACATGGTCTTAGCGTATGACCTTGCCTTAGGGAGTGGCTT 1140  
QY 381 ProLeuAsnArgSerValAsnAsnValValAlaArgTyrGlyHisSerLeuAlaLeuTyr 400  
DB 1141 CCACATAAACCCGTCTGTGAAACATGTGGTGTGTAGATATGGTCAATCTTGGCATATATAC 1200  
QY 401 LysAspLysIleTyrMetTyrGlyLysIleAspProThrGlyAsnValThrAsnGlu 420  
DB 1201 AAGGATAAATTTACATGTATCGAGGAATAATGATCCACTGGGAATGTGACCAATGAG 1260  
QY 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440  
DB 1261 TTTCAGAGTTTTTCACATTCATAATGAGTCATGGGTGTTGTTGACCCCTTAAAGCAAGGAG 1320  
QY 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460  
DB 1321 CAGTATGCAGTGGTGGGCACCTCTGCACACATGTGTACACTGAAGATGGCGAGTGGTC 1380  
QY 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480  
DB 1381 ATGCTGGTCACTCTTTGGTCACTGCCCTCTCTATGATATATAAGCAATGTGCGAGGAATAT 1440  
QY 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500  
DB 1441 GATTTGGATPAGAACAACATCGGAGTATATTACACACCCAGGGTGGCCCTTGTGCAAGGGGT 1500  
QY 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
DB 1501 TAGGCCATAGCAGTGTTTACAGCCATAGGACACAGGGCCCTATACGTTCTAGGTGGCTAC 1560  
QY 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540  
DB 1561 AAGGGTTTTCAGTGGCCAAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620  
QY 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560  
DB 1621 ACCCAGATGTGGACCATCTTTAAGGACAGCCGATTTTCCCGTTACTTGCACACAGCTGTG 1680

QY 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
DB 1681 ATAGTGAGTGAACCATGCTGGTGTGTTGGGGAAACACACACAATGACATCATGAGC 1740  
QY 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600  
DB 1741 CATGGCGCAATGCTTCTCTTCAGATTTTCATGSCCTATGACATTTGCTGTGACCGCTGG 1800  
QY 601 SerValLeuProArgProAspLeuHisAspValAsnArgPheGlyHisSerAlaVal 620  
DB 1801 TCAGTGTCTCCAGACCATGATCTCCACCATGATCAACAGATTTGGCCATTCAGCAGTC 1860  
QY 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuLeuSerAspIle 640  
DB 1861 TTACACAACAGCACCATTGATGTTTGGTGGTTCATAGTCTCTCTCCTCAGCGACATC 1920  
QY 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660  
DB 1921 CTGGTATTTCCTCGAACAGTGTGATCGCATCGAGTGAAGCCGCTTGTGTTAGCAGCA 1980  
QY 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680  
DB 1981 GGACTCTGTTATTCGTTGTGTGTGGAACACAGGGTCTCTCAGTGTATCTCGTGGGCCCTG 2040  
QY 681 AlaThrAspGluGlnGluGlyLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700  
DB 2041 GCAACTGATGAACAAGAAAGAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTTGAC 2100  
QY 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720  
DB 2101 CATGACAGATGTGACCCAGCACACAGATTTGTACAGCTGTACAGCCAAACCAATGACTGC 2160  
QY 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
DB 2161 CACTGGTGCAATGACCATTTGTGTCGCCAGGAACACAGCTGCTCAGAGGCCAGATCTCC 2220  
QY 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760  
DB 2221 ATTTTATAGTATGAGAAATTCGCCCAAGSATAACCCCATGTACTACTGTAAACAAGAAC 2280  
QY 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780  
DB 2281 AGCTGAGGAGCTGTGCCCTGGACAGAACTGCCAGTGGGAGGCCCGGATCAGGAGTGC 2340  
QY 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800  
DB 2341 ATTGCCCTGCCGAAAATATCTGTGGCATTTGGTGGCATTTGGTGGAACTCATGTTTG 2400  
QY 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820  
DB 2401 AAAATTTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGAACCAAT 2460  
QY 821 AlaLeuLeuAlaSerLeuThrThrGlnLysValGluPheValLeuLysGlnLeuArg 840  
DB 2461 GCCCTTTTGGCTTCTCTTACACCCAGAGAAGTAGAATTTGTCCTTAAGCAGCTGCGA 2520  
QY 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860  
DB 2521 ATAATGCACTCATCTCAGACATGTCCAGCTCACCTTAACCCCATGGGTGGGCTTCGG 2580  
QY 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880  
DB 2581 AAGATCAATGTGCTTACTGCTGGAGAGATATGTCCCCATTTACAAATAGTTTACTA 2640  
QY 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
DB 2641 CAGTGGATGCGCTGTGAGCCAGTGTCTGTGGATTTCTGTGGAATTTTATCAGAACCCAGT 2700  
QY 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
DB 2701 ACTGGGGATGAAGGCTGCACTGCATCAACCCACTCAATGATGTGTGTGTGAAAGG 2760  
QY 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940







QY 597 CysAspArgTyrSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616  
Db 1786 TGTGACCGCTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845  
QY 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
Db 1846 CATTACAGCAGTCTTACCAACAGCACCATGTATGTGTGGTGGTTCAATAGTCTCCTC 1905  
QY 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
Db 1906 CTCAGCGACATCTGTGTAATTCACCTCGGAACAGTGTGATCGCATCGGAGTGAAGCCGCT 1965  
QY 657 CysLeuAlaAlaGlyProGlyIleArgCysValTyrAsnThrGlySerSerGlnCysIle 676  
Db 1966 TGTATTAGCAGCAGGACCTGGTATTCGTGTGTGTGTGGAACACAGGTCGTCTCAGTGTATC 2025  
QY 677 SerTyrAlaLeuAlaThrAspGluGlnGluGlyLeuLeuSerGluCysPheSerLys 696  
Db 2026 TCGTGGCGCTGGCACTGATGAACAGAGAAAGTTAAATCAGATGTTTTTCCAAA 2085  
QY 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
Db 2086 AGAACTCTTGACCATGACAGATGTGACCAAGCACACAGATTTGTACAGCTGCACGCCAAC 2145  
QY 717 ThrAsnAspCysHisTyrCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
Db 2146 ACCAATGATGCCACCTGCTGTCATGACCAATTTGTCTCCAGGACACACAGCTGCTCAGAA 2205  
QY 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
Db 2206 GGCCAGATCTCCATTTTATAGTATGAGAATTCGCCCAAGGATAACCTATGTACTACTGT 2265  
QY 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTyrProArg 776  
Db 2266 AACAAAGAACAGCTGCGAGAGCTGTGCCCTGGACAGCAATGTCAGTGGAGGCCCGG 2325  
QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTyrHisLeuValGly 796  
Db 2326 AATCAGAGTGCAATGCCCTGCGCGAAATATCTGTGGCAATTCGTGGCATTTGGTTGGA 2385  
QY 797 AsnSerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
Db 2386 AACTCATGTTTGAATAATTAATCTGCAAGGAGAAATATGACAAATCTAAATGTTCCT 2445  
QY 817 ArgAsnHisAsnAlaLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836  
Db 2446 AGAAACCAATGCGCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGTCCTT 2505  
QY 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTyr 856  
Db 2506 AAGCAGCTGCGAATAATGCACTCATCTCAGAGCATGTCCAAGCTCACTTAACCCCATGG 2565  
QY 857 ValGlyLeuArgLysIleAsnValSerTyrTyrCysTyrGluAspMetSerProPheThr 876  
Db 2566 GTCGGCTTCGGAAGATCAATGTCTCTACTGTGTGTGGAGATATGTCCTCATTTTACA 2625  
QY 877 AsnSerLeuLeuGlnTyrMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
Db 2626 AATAGTTTACTACAGTGGATGCGCTCTGAGCCAGTGTGTCGATTCGTGGAAATTTTA 2685  
QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
Db 2686 TCAGAACCCAGTACTCGGGACCTGAAGGCTGCAACTGCAATCAACCATCATCATGTAGT 2745  
QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
Db 2746 GTCTGTGAAGGCCCTGCAACACCAAGTGTAAAGCAGTGCAGGACACCATGTGCTTGGG 2805  
QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLys 956  
Db 2806 ACAGCATGTGAGATTCACACAGCGGAGCTCTGAGTGCATGTCTGAGTGCAGCAATGAG 2865

QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyr 976  
Db 2866 CAGTGTGTGACTCCATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGG 2925  
QY 977 TyrThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
Db 2926 TATACGATGAGCACCTGCCCCCTGAAATTTGTTACGGCTACTGTACCTGTGTAGTCAATGC 2985  
QY 997 LeuGluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCys 1016  
Db 2986 TTGAGCAACACGCTGTGGTGTGTACTGATCCAGCAATACTGCGAAAGGGAAATGC 3045  
QY 1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036  
Db 3046 ATAGAGGCTTCTTATAAAGGACCACTGAAGATGCTTCGCAAGCCCTACAGGAATTTTC 3105  
QY 1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSer 1056  
Db 3106 TATCCAGACCCCTGCTCAATTCAGCATGTGCTTAGAGGACAGCAGATACAACTGGTCT 3165  
QY 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
Db 3166 TTCATTCACTGTCCAGCTTCCCAATGCACCGGCACAGTAAATGTCATCAATCAGAGCATC 3225  
QY 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096  
Db 3226 TGTGAGAAGTGTGAGAACCTGACACAGGCAAGCCTGCGAGACCTGCATATCTGCGTTC 3285  
QY 1097 TyrGlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116  
Db 3286 TACGGTCAATCCCAATGGAGGGAAATGTACGCCATGCAAGTGCATGGCACCGGCTCT 3345  
QY 1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136  
Db 3346 CTGTGCAACCAACACACGGCAAGTGTCTCTGCACCAAGGGCGTCAAGGGGGACGAG 3405  
QY 1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156  
Db 3406 TGCAGCTATGTGAGGTAGAATTCGATACCAAGGAACCTCTCAGAGGAACATGTTAT 3465  
QY 1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnLeuAspAspArgTyr 1176  
Db 3466 TATACTCTTCTTATTGACTATCAGTTTCACTTACCTTCTATCCAGGAAGATGATCGCTAT 3525  
QY 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196  
Db 3526 TACACAGCTATCAATTTTGGCTTCTCTGACGACAAACAAACAGGGATTTGGACATGTTTC 3585  
QY 1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216  
Db 3586 ATCAATGCCCTCAAGAATTTCACTCAACATCACCTGGCTGGCTGCCAGTTCTCAGCTGGA 3645  
QY 1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236  
Db 3646 ACCCAGGCTGGAGAAGAGATGCTCTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705  
QY 1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256  
Db 3706 AGTTTCTTAATGAGAAGTTTGAATTTTCGCAACCCACCAATATCACTTCTTGTGTTAT 3765  
QY 1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276  
Db 3766 GTCAAGTAAATTTTCACTGGCCCATCAAAATTCAGATTCCTTCTCAGCAGCAAGCAATTTT 3825  
QY 1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuVal 1296  
Db 3826 ATGACCTGTGTACAGTCTTCTGTCATCTTCTTCAAGTTGTTTCTCTCTTCTCTGCTG 3885  
QY 1297 AlaAlaValValTrpLysIleLysGlnSerCysThrAlaSerArgArgGluGlnLeu 1316  
Db 3886 GCTGTGTGTGTGAAGATCAAAACAAAGTTGTGGGCTCCAGACCTAGAGAGCAACTT 3945  
QY 1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336



Db 3946 CTTGAGAGATGCAACAGATGGCCAGCGTCCCTTTGCTCTGTAATGTCGCTTGGAA 4005  
 QY 1337 ThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356  
 Db 4006 ACAGATGAGAGAGCTCTGATCTTATTGGGGGAGTATAAAGCTGTTCCTCCAAACCCATT 4065  
 QY 1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376  
 Db 4066 GCACTGGAGCGGTGTTTGGCAACAAAGCGGTCTCTCTGTGTGTGTGAGGCTCCCT 4125  
 QY 1377 ArgGlyLeuGlyGlyLeuProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396  
 Db 4126 CGAGGCTGGTGGCAATCCCTCTCTCTGGGCACTAGGTCCTGTGTGGCCAGCCCTTG 4185  
 QY 1397 ValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsn 1416  
 Db 4186 GTGGACATTCTCAGCAGATGCGATAGTGTACAGGAGAAGTCAGGAGCCGTGAGAAAC 4245  
 QY 1417 ArgLysGlnProProGlnProGlnProGlnProGlyThrCysIle 1429  
 Db 4246 CGAAGCAGAGCCCTGCACAGCTGGGACCTGCATC 4284

## RESULT 4

AAZ52275

ID AAZ52275 standard; cDNA; 4068 BP.

XX AC AAZ52275;

XX DT 18-JUL-2000 (first entry)

XX DE Human membrane attractin-1 cDNA.

XX KW Human; membrane attractin-1; immune response; macrophage; monocyte;

XX KW T cell; immunostimulant; immunosuppressed patient; cancer;

XX KW immunodeficiency syndrome; transplant; autoimmune disease; ss.

XX OS Homo sapiens.

XX FH Key

XX FT CDS

XX FT Location/Qualifiers

XX FT 1..4068

XX FT /tag= a

XX FT /product= "Membrane attractin-1"

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

XX FT

CC patients such as those undergoing chemo- and radio-therapy treatment for  
 CC cancer or those suffering from common variable immunodeficiency syndrome.  
 CC The proteins may also be used to screen modulators (agonists and  
 CC antagonists) of immune response which may also be used to regulate immune  
 CC reactions. Attractin antibodies can be used to inhibit immune response in  
 CC transplant recipients or patients afflicted with autoimmune disease  
 XX  
 SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0 Length: 4068  
 Score: 7479.00 Matches: 1353  
 Percent Similarity: 94.75% Conservative: 1  
 Best Local Similarity: 94.68% Mismatches: 1  
 Query Match: 94.48% Indels: 74  
 DB: 3 Gaps: 1

US-09-787-097-12 (1-1429) x AAZ52275 (1-4068)

QY 1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAla 20  
 Db 1 ATGTTGGCCGCGCGGCGCACTGAGGCAAGGCTGAGGAGGAGCGCGCGACGGCA 60  
 QY 21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpValThrArgAlaGlyArg 40  
 Db 61 GCGTCTCGGGGAGGAGCGGGCGCCGAC----- 90  
 QY 41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60  
 Db 90 ----- 90  
 QY 61 LeuLeuLeuLeuLeuLeuLeuProProLeuLeuLeuLeuLeuProCysGlu 80  
 Db 90 ----- 90  
 QY 81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100  
 Db 90 ----- 90  
 QY 101 CysAspArgProCysValAsnGlyArgCysAsnProGlyThrGlyGlnCysValCys 120  
 Db 91 -----TGTGTCAACGGCGTCTGCTGCAACCTGCGACCGCGCCAGTCGCTGTC 138  
 QY 121 ProAlaGlyTrpValGlyGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140  
 Db 139 CCGCGCGCTGGTGGCGGAGCAATGCCAGACATGCGGGGGCGCGCTCAGACTAATGGA 198  
 QY 141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160  
 Db 199 TCTTCTGGGTTTGTGACAGATGGACCTGGAAATTTATAAATACAAAACGAAGTGCACGTGG 258  
 QY 161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180  
 Db 259 CTCATTGAAGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCAATTTTGTACAGAG 318  
 QY 181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200  
 Db 319 TGTAGTTGGGACCAATTATATGTTTATGATGGGACCTCAATTTATGACCGCTAGTTGCT 378  
 QY 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220  
 Db 379 GCATTTAGTGGCTTCATTGTTCTCCAGAGAGATGGCAATGAGACTGTCCCTGAGGTGTT 438  
 QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240  
 Db 439 GCCACATCAGTTATGCTTGTCTGCTGCAATTTTATGATGCTGCTTATATTTGACTGGA 498  
 QY 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260  
 Db 499 TTTAATATTACTTACAGTTTGTATGTGTCCAAATAACTGCCTCAGGCCGAGGAGAGTGT 558  
 QY 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280

Isolated nucleic acids encoding human attractin polypeptides useful for enhancing immune responses.

Claim 3; Fig 10; 120pp; English.

The patent discloses four forms of human attractin polypeptides which enhance immune response by promoting macrophage and monocyte spreading in the presence of T cells. These include soluble attractin-1 and -2 and membrane attractin-1 and -2. These various forms of attractin are encoded by alternatively spliced mRNA molecule transcribed from a single gene. The present sequence is a cDNA encoding human membrane attractin-1. Membrane attractin differs from soluble attractin in having a transmembrane domain and a cytoplasmic domain at the C-terminal. This sequence can be used to enhance immune response in immunosuppressed

559	Db	AA	GATCAGTAATAAGCAGCGGCTACTGTTGAATGTGAATGTTCTCGAAACTCGAAAGGTGAA	618
281	Qy	Ala	CysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn	300
619	Db	GA	TGTGACATTCCTCCTCCTGACAGCAACTGTGGTTTCTCTCATCGAGCAATCGCAAT	678
301	Qy	Ser	SerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal	320
679	Db	TCA	AGTCATGTCCAGAGGATGCTCTGCTTCTCAGACTGCGCAGGCGTCTCGATGTTCCAGTT	738
321	Qy	Pro	ValProAlaAsnGlnSerPheTrpThrArgGlyGluTyrSerAsnLeuLysLeuPro	340
739	Db	CC	TGTACCAGCTAACAGTCATTTTGACATCGAGAGGATATTTCTTAACCTTAAGCTCCCC	798
341	Qy	Arg	AlaSerHisLysAlaValAlaAsnGlyAsnIleMetTrpValValGlyGlyTyrMet	360
799	Db	AG	AGCATCTCATAAAGCTGTGGTCAATGGGAAACATTATGTGGGTGTGAGGATATATG	858
361	Qy	Phe	AsnHisSerAspTyrAsnMetValLeuAlaTyrAspIleuAlaSerArgGluTrpLeu	380
859	Db	TT	CAACACCTCAGATTATAACATGGTGTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTT	918
381	Qy	Pro	LeuAsnArgSerValAsnAsnValValAlaArgTyrGlyHisSerLeuAlaLeuTyr	400
919	Db	CC	ACTAAACGTTCTGTGAACAATGTGGTGTGTAGATATGTCATCTCTTTGGCATTATAC	978
401	Qy	Lys	AspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu	420
979	Db	AAG	GATAAAATTTACATGTATGAGGAGAAATTTGATTTCAACTGGGATGTGACCAATGAG	1038
421	Qy	Leu	ArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu	440
1039	Db	TT	GAGGTTTTCCATTCATATCACTCATATCACTCATGGGTGTGTGTGACCCCTAAGGCCAAGGAG	1098
441	Qy	Gln	TyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal	460
1099	Db	CAG	TATCGACTGGTTGGGCACTCTGACACACATTTGTTACACTGAAGAATGCCGAGTGGTC	1158
461	Qy	Met	LeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr	480
1159	Db	AT	CGTGTGCTCATCTTTGTGTCACTGCCCTCTCTATGGATATATAGCAATGTGTCAGGAATAT	1218
481	Qy	Asp	LeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly	500
1219	Db	GA	TTTGGATTAAGAACACATCGAGTATATTACACCCAGGGTGGCTTGTGCAAGGGCGT	1278
501	Qy	Tyr	GlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr	520
1279	Db	TAC	GGCCATAGCAGTGTATTACACCATAGACACAGGGCCCTATACGTTTCATGGTGGCTAC	1338
521	Qy	Lys	AlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrAArgTyrAspValAsp	540
1339	Db	AAG	CTTTCAGTGCCATATAGTACCGCTTCGACATGATCTCTACCGATATGATGCGAT	1398
541	Qy	Thr	GlnMetTrpThrIleLeuLysAspSerArgPheArgTyrLeuHisThrAlaVal	560
1399	Db	ACC	CAGATGTGGACCATTTCTTAAGGACAGCGCGATTTTTTCGGTTACTTTGCAACAGCTGTG	1458
561	Qy	Ile	ValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer	580
1459	Db	ATA	GTTAGTGAACCATGCTGGTGTTTGGGGGAAACACACACATGACACATCTATGAGC	1518
581	Qy	His	GlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp	600
1519	Db	CAT	GGCGCCAAATGCTCTCTCTTCAGATTTTCATGGCCCTATGACATTCGCTGTGACCGGTGG	1578
601	Qy	Ser	ValLeuProArgProAspLeuHisIleAspValAsnArgPheGlyHisSerAlaVal	620
1579	Db	TC	AGTGTCTCCACACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTTCAGCACTC	1638
621	Qy	Leu	HisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIle	640
1639	Db	TT	CACAAACAGCACCATGTATGTTCGGTGGTTTCAATAGTCTCTTCCTCCTCAGCAGATC	1698

Qy	641	LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla	660
Db	1699	CTGGTATTACCTCGGAACAGTGTGTATGGCATCGAGATGAAGCCGCTGTGTAGCAGCA	1758
Qy	661	GlyProGlyIleArgCysValTtpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu	680
Db	1759	GGACCTGGTATTCCGGTGTGTGGAAACAAGGTGCTCTCAGTGTATCTCGTGGGCGCTG	1818
Qy	681	AlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp	700
Db	1819	GCAACTGATGAACAAGAGAAAGTAAAAATCAGAATGTTTTTCCAAAAACAATCTTTGAC	1878
Qy	701	HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys	720
Db	1879	CATGCAGAGTGTACCAAGACACACAGATTGTTACAGCTGCACAGCAACCAATGACTGC	1938
Qy	721	HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer	740
Db	1939	CACCTGGTCAATGCCATTGTGTCCCCAGAAACCAAGCTGCTCAGAAAGCCAGACTCC	1998
Qy	741	IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr	760
Db	1999	ATTTTTAGGTATGAGAATTGCCCCAAGATAAACCCTATGTACTACTGTATAACAAGAAGACC	2058
Qy	761	SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys	780
Db	2059	AGCTGCAGGAGCTGTGCCCTGGACCAAGACTGCCAGCTGGAGGCCCGGAATCAGGAGTGC	2118
Qy	781	IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyIleAsnSerCysLeu	800
Db	2119	ATTGCCCTGCCGAAATAATCTCTGCATTTGGCTGGCATTTGGTTGGAACTCATGTTTG	2178
Qy	801	LysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn	820
Db	2179	AAAATTACTACTGCCAAGAGAAATTATGCAATGCTTAATTGTTCTGTAGAACCAACAT	2238
Qy	821	AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg	840
Db	2239	GCCCTTTTGGCTTCTTTACAACCCAGAAAGGTAGAAATTTGCTCTTAAGCAGCTGCCA	2298
Qy	841	IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg	860
Db	2299	ATAATGCAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGGTCGGCTTCGG	2358
Qy	861	LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu	880
Db	2359	AAGATCAATGTGTCTTACTGTGCTGGAGATATGTCCCCCAITTAACAATAGTTTACTA	2418
Qy	881	GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer	900
Db	2419	CAGTGGATGCGTCTGAGCCAGTGTGCTGTGGATTTCTGTGAAATTTTATCAGAACCCAGT	2478
Qy	901	ThrArgGlyLeuLysAlaIleThrCysIleAsnProLeuAsnGlySerValCysGluArg	920
Db	2479	ACTCGGGATCTGAAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGG	2538
Qy	921	ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly	940
Db	2539	CCTGCAAAACCAAGTGTAAAGCAGTGCAGACACCAATGTGCCCTTGGAGACAGCATGTGGA	2598
Qy	941	AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp	960
Db	2599	GATTGCACCAAGCGCAGCTCTGAGTGTGCATGTGGTGCAGCAACATGAACGAGTGTGTGCAC	2658
Qy	961	SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer	980
Db	2659	TCCAAATGCCATGTGGCCCTCCTTCCTTTTGGCCAGTGTATGGAAATGGTATACGATGAGC	2718
Qy	981	ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro	1000
Db	2719	ACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCATTCCTCTGGAGCAACCA	2778





Db 1861 GACCGATGGTCAGTCTCTCCACAGCTGAGCTCCATCAGATGTCAACCGATTGGCCAC 1920  
Qy 618 SerAlaValLeuHisSerThrMetTyrValPheGlyPheAsnSerLeuLeuLeu 637  
Db 1921 TCAGCAGTCTTGACAAACAGCACCATGTATGTGTTCGGGGCTTCAACAGCCCTCTCTC 1980  
Qy 638 SerAspIleuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaCys 657  
Db 1981 AGTGACGTCTTAGCTTTCACCTCGGAGCAGTGTGATGCACACCGCAGGAGCTCTGT 2040  
Qy 658 LeuAlaAlaGlyProGlyIleArgCysValTyrAsnThrGlySerSerCysIleSer 677  
Db 2041 GTGGCAGCAGGACCTGGCATCGGTCTCTGTGGGACACACAGCTGTCTGATGTACCTCC 2100  
Qy 678 TrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerIysArg 697  
Db 2101 TGGGAGTGGCACTCGAAGAACAGCAGAAAGTTAAATCAGAATGTTTCTCTAAAAGA 2160  
Qy 698 ThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThr 717  
Db 2161 ACCCTTGACATGACAGATGTGACACAGCACACAGATTGTACAGCTGCACAGCCATACC 2220  
Qy 718 AsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGly 737  
Db 2221 AATGACTGCACATGGTGTCAATGATCACTGTGTCCCTGTGAACACACAGCTGCACAGAAGC 2280  
Qy 738 GlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsn 757  
Db 2281 CAGATCTCCATTGCCAAGTATGCAATGCCAGAGTAACCCATGACTACTACTGCAAT 2340  
Qy 758 LysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsn 777  
Db 2341 AAGAAACACAGCTGCAGGAGCTGTGCCCTAGATCAGAAATTGCCAGTGGGAACCTCGAAAT 2400  
Qy 778 GlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsn 797  
Db 2401 CAAGAGTGCATGCCCTGCCGAAACATCTGTGGCATTGGCTGGCATTGGTCCGGAAC 2460  
Qy 798 SerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArg 817  
Db 2461 TCGTGTCTGAATCACTACTCTAAGGAAATATGACAAATGCTAATGTCTCTGTAGA 2520  
Qy 818 AsnHisAsnAlaLeuLeuAlaSerLeuThrGlnLysLysValGluPheValLeuLys 837  
Db 2521 AACCACAAATGCCTTTTGGCTTCGCTTAGCTCCAGAACAGTGTGGCATTGGTCCGGAAC 2580  
Qy 838 GlnLeuArgIleMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTrpVal 857  
Db 2581 CAGCTTCGATTAATGCATCACTTCAGACACGTCCAGCTCACTGACTCCATGGGTT 2640  
Qy 858 GlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsn 877  
Db 2641 GGTCTTCGAAGATCAATGTATCTTACTGTGTGGGAGATATGTCTCCATTCAAAAT 2700  
Qy 878 SerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSer 897  
Db 2701 AGTTGTCTCAGTGGATGCCATCTGAGCCAGTGTAGTCCGGCTCTGTGGGATCTGTCA 2760  
Qy 898 GluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerVal 917  
Db 2761 GAGCCTAGTACTCGGGACTGAGGCTGCAACTGCATCAACCCACTCAATGGAAGTGT 2820  
Qy 918 CysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThr 937  
Db 2821 TGTGAAGCCCTGCAACACACAGTGCACAGCAGTCCGACACCCCTGTGCCCTGGCGACA 2880  
Qy 938 AlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGln 957  
Db 2881 GCATGTGTGTGTCACCGACAGCAGCTCCGATGATGTGTGTGTGTGTGTGTGTGTGTGT 2940  
Qy 958 CysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyr 977  
Db 2941 TGGGTAGATCCAAATGCCTACGTAGCTCTCTTTCCTTTTCGGCCAGTGTATGGAGTGTAT 3000

Qy 978 ThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
Db 3001 ACAATGAGCAGCTGCCACCTGAAATTTGCTCTGGCTACTGTACTCTGAGCCATTGCTTG 3060  
Qy 998 GluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIle 1017  
Db 3061 GAGCAGCAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3120  
Qy 1018 GluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
Db 3121 GAGGGAAGCTATAAAGGACCCGTGAAGTGCCTTCACATGCTCTACAGGAATGTGTAC 3180  
Qy 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057  
Db 3181 CCACAGCCCTTTCGAACTCCAGCATGTGTCTAGAGACACAGATACAACCTGGCTCTTT 3240  
Qy 1058 IleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCys 1077  
Db 3241 ATTCACTGTCCAGCTTGCAGTGCATGCAATGGACACAGCAAGTGCATCAACAGAGCATCTGT 3300  
Qy 1078 GluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyr 1097  
Db 3301 GAGAAGTGTGAGGACCTGCACACAGGCAAGCACTTCGAGACCTGCATATCTGGCTCTAT 3360  
Qy 1098 GluAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117  
Db 3361 GGTGACCCGACTAATGAGGCAATGTACCCATGCAAGTGCACACGGGACGCACTACTG 3420  
Qy 1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137  
Db 3421 TGCAACCAACACTGGCAAGTGTCTGTACCAACCAAGGCGCTCAAGGGAGAGAGAGTGC 3480  
Qy 1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157  
Db 3481 CAGTATGTGAGGTAGAAAATCGATACAGGGAACCCCTCTCAAGGAACATGTTACTAT 3540  
Qy 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177  
Db 3541 ACCCTTCTCATTGACTATCAGTTTCACTTTAGCTGTCCCAGGAAGACGACCGCTATTAC 3600  
Qy 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197  
Db 3601 ACAGCCATCAATTTTGGCTACTCTCTGTAGAACAAACAGGGATTTGGACATGTTTCATC 3660  
Qy 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThr 1217  
Db 3661 AATGCTTCCAAAAAATTTCAACCTCAACATCACCTGGGCCACAGTTTCCAGCTGGAAAC 3720  
Qy 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237  
Db 3721 CAGACTGGAGAGAGGTGCCTGTGTTTCAAAAAACCAACATCAAGGAATACAAAGATAGC 3780  
Qy 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257  
Db 3781 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATTAATTTCTTTGTTTATGTC 3840  
Qy 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277  
Db 3841 AGTAATTTCACTTGGCCCATCAAAATTCAAATTCGCTTCTCCAGCACACAGCAATTTATG 3900  
Qy 1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAla 1297  
Db 3901 GACCTGTGAGTCTTTCGTGACTTTCTTTCAGCTGTTTCTCTCTCTCTCTCTCTCTCT 3960  
Qy 1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgAspGluGlnLeuLeu 1317  
Db 3961 GCAGTGTCTGGAGATCAAGCAGCTGTTGGGCATCCAGGAGGAGAGAGCACTTCTT 4020  
Qy 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337  
Db 4021 CGGAGATGCAACAGATGGCCAGCCGCTTTTCTTCTGTAATGTTGCTTCTTCTTCTTCTT 4080









Db	3059	ACCATGTCGCCGACAGCGTGTGGCGAGTGCATACGACGAGCTCGGAGTCATGTG	3118
Qy	951	pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPheProPheG1	971
Db	3119	GTGCAGTAACATCAAGACAGTGTGTGACATCCCAATGCTACGTGGCTCTCCCTTTTGG	3178
Qy	971	yGlnCysMetGluTyrThrMetSerThrCysProGluAsnCysSerGlyTyrCy	991
Db	3179	CCAGTGTATGGGAATGTATACGATGAGCAGCTGCCCACTCGAATATGCTCTGGGTATG	3238
Qy	991	sThrCysSerHisCysLeuGluGlnProGlyCysGlyTyrCysThrAspProSerAsnTh	1011
Db	3239	TACCTCGACCATTTGCTTGGAGCAGCAGCGCTGTGGTGTCTACTGATCTCTAGCAATAC	3298
Qy	1011	rglyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAl	1031
Db	3299	TGGGAAGGAAATGTATTGAGGGCAGCTATAAGAGACCTGTGAAGATGCCGTACAGGC	3359
Qy	1031	aProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSe	1051
Db	3359	CTCTGCAGGAATGTGTATCCACAGCCCTTCTGAACCTCCAGCATGTGTCTAGAGGACAG	3418
Qy	1051	rArgTyrAsnTyrSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCy	1071
Db	3419	CAGATCAACTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	3478
Qy	1071	sIleAsnGlnSerIleCysGluLysCysGluLeuLeuThrThrGlyLysHisCysGluTh	1091
Db	3479	CATCAACCCAGAGTATCTGTAGAGTGTGTGAGACCTGACCAAGGCGACCATGCGAGAC	3538
Qy	1091	rCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCy	1111
Db	3539	CTGCATATCTGGCTTCTTATGGTACCCGACTATATGGAGGCAAAATGTTCAGCCATGCA	3598
Qy	1111	sAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysG1	1131
Db	3599	CAATGGCAGCATCATCTGTGCAACACCAACACCGGCAAGTCTTCTGTACCAACCAAGG	3658
Qy	1131	yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLe	1151
Db	3659	TGTCNAGGGGACGAGTGGCAGCTATGTGAGGTAGAAATCGATACCAAGGAAACCTCT	3718
Qy	1151	uArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerG1	1171
Db	3719	CAAGGAAACATGTCTACTATACCTTCTCATTTGACTATCAGTTACCTTTAGCCTGTCCCA	3778
Qy	1171	nGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnAr	1191
Db	3779	GGAGACGACCGCTACTACACGCCATCACTTTTGTGGTACTCTCTGTATGAACAAACAG	3838
Qy	1191	gAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAl	1211
Db	3839	GGATTTGGACATGTTCAATCAATGCCTCCAAAACCTTCAACCTCAACATCACCTGGGCCAC	3898
Qy	1211	aSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnI1	1231
Db	3899	CAGCTTCCCGACCGGAACCCAGACTGGAGAGAGGTGCTGTGTGTTTCAAAACCAACAT	3958
Qy	1231	eLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnI1	1251
Db	3959	CAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACGATCCAAACAT	4018
Qy	1251	eThrPhePheValTyrValSerAsnPheThrTyrProIleLysIleGlnIleAlaPheSe	1271
Db	4019	CACTTTCTTTTGTATGTCTAGTAATTTCACTTGGCCCATCAAAATTCAGATTCCTTCTC	4078
Qy	1271	rGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLe	1291
Db	4079	CCAGCACAGCAACTTCATGACCTGTGTACAGTTTCTTCGTGACTTTCTTCAGTGTGTCT	4138
Qy	1291	uSerLeuLeuValAlaAlaValTyrLysIleLysGlnSerCysThrAlaSerAr	1311
Db	4139	CTCGCTCTTCTGTGGTGCAGTGTGTCTGGAAGATCAACGACAGCTTTTGGGCATCCAG	4198
Qy	1311	gArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerVa	1331
Db	4199	CGCGAGAGAGCAACTTCTTTCGGAGATGCAACAGATGGCCAGCGCCCTTTGCTTCTGT	4258
Qy	1331	lAsnValAlaLeuGluThrAspGluGluProProAspLeuIleGlyLysIleLysTh	1351
Db	4259	AAACGTTTGCCTTGGAAACAGATGAGGAGCTCTCTATCTTATTGGGGGGAGTATAAGAC	4318
Qy	1351	rValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerVa	1371
Db	4319	TGTTTCCCAACCATTTGACTGTGAGCGGTGTTTGGCAACAGCCGCTGTCTCTCTGT	4378
Qy	1371	lPheValArgLeuProArgGlyLeuGlyLysIleProProGlyGlnSerGlyLeuAl	1391
Db	4379	GTTTGTGAGGCTCCCTCGAGCCCTGGGTGGCATCCCTCTCTCTGGCAGTCAAGTCTTGC	4438
Qy	1391	aValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSe	1411
Db	4439	TGTGGCCAGCGCCCTGTGTGACATTTCTCAGCAGATGCCGATAGTGTACAGGAGAGTC	4498
Qy	1411	rGlyValAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle	1429
Db	4499	AGGAGCCGTGAGAAACCGGAAGCAGCAGCGCCCTGCACAGCCTGGGACCTGCATC	4553
RESULT 7			
ID	AAZ52277	standard; cDNA; 3819 BP.	
XX	AAZ52277		
AC	AAZ52277;		
DT	18-JUL-2000	(first entry)	
DE	Human soluble attractin-2 cDNA.		
XX	Human; soluble attractin-2; immune response; macrophage; monocyte;		
KW	T cell; immunostimulant; immunosuppressed patient; cancer;		
KW	immunodeficiency syndrome; transplant; autoimmune disease; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	1..3819	
FT		/*tag= a	
FT		/product= "Soluble attractin-2"	
XX	WO200015651-A1.		
PN	23-MAR-2000.		
PD	14-SEP-1999;	99WO-US020948.	
PF	14-SEP-1998;	98US-0100137P.	
PR	(DAND )	DANA FARBER CANCER INST INC.	
XX	Duke-Cohan JS,	Schlossman SF;	
PA	WPI;	2000-271373/23.	
PI	P-PSDB;	AA70692.	
XX	Isolated nucleic acids encoding human attractin polypeptides useful for		
PT	enhancing immune responses.		
PS	Claim 3; Fig 12;	120pp; English.	
XX	The patent discloses four forms of human attractin polypeptides which		
CC	enhance immune response by promoting macrophage and monocyte spreading in		
CC	the presence of T cells. These include soluble attractin-1 and -2 and		
CC	membrane attractin-1 and -2. These various forms of attractin are encoded		
CC	by alternatively spliced mRNA molecule transcribed from a single gene.		
CC	The present sequence is a cDNA encoding human soluble attractin-2.		
CC	Attractin-2 differs from attractin-1 in having a 74 amino acid insert in		

CC the N-terminal. This sequence can be used to enhance immune response in  
 CC immunosuppressed patients such as those undergoing chemo- and radio-  
 CC therapy treatment for cancer or those suffering from common variable  
 CC immunodeficiency syndrome. The proteins may also be used to screen  
 CC modulators (agonists and antagonists) of immune response which may also  
 CC be used to regulate immune reactions. Atractin antibodies can be used to  
 CC inhibit immune response in transplant recipients or patients afflicted  
 CC with autoimmune disease  
 XX

SQ Sequence 3819 Bp; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0 Length: 3819  
 Score: 7099.00 Matches: 1268  
 Percent Similarity: 99.76% Conservative: 1  
 Best Local Similarity: 99.69% Mismatches: 3  
 Query Match: 89.56% Indels: 0  
 DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x AAZ52277 (1-3819)

QY 1 MetValAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAla 20  
 Db 1 ATGGTGGCGCGAGCGCGCACTGAGGCAAGGCTGAGGAGGACGGCGCGCGCA 60  
 QY 21 AlaLeuAlaGlyArgSerGlyProHisTTPAspTTPAspValThrArgIadIyArg 40  
 Db 61 GGCTCGCGGGGAGGAGCGCGCGCACTGGGACTGGGAGCGTGACCGAGGCTGGGAG 120  
 QY 41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60  
 Db 121 CGGGGCTGGGCGCGCGTGGCTCCGCGGCTGCTCTCCACCGTGGCGCACGG 180  
 QY 61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80  
 Db 181 CTGCTCTGCTGCTGTGTGTGCTCCGCGCGCGCTGTGCTGCTGCTGCTGCTGAG 240  
 QY 81 AlaGluAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100  
 Db 241 GCGGAGCGCGCGCGCGCGCGGCGGTGCGGCTCAGCGCGAGCGGAGCGCAAGGA 300  
 QY 101 CysAspArgProCysValAsnGlyArgCysAsnProGlyThrGlyGlnCysValCys 120  
 Db 301 TGTGACCGCGCTGTGTCAACGCGCGTGTGCTGCAACCGCTGGCACCGCGCTGCTGC 360  
 QY 121 ProAlaGlyTTPValGlyGlnCysGlnHisCysGlyArgPheArgLeuThrGly 140  
 Db 361 CCGCGCGCTGGTGGCGGAGCAATGCCAGCACTGGCGGCGCGCTTCAGACTAACTGGA 420  
 QY 141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTTP 160  
 Db 421 TCTTCTGGTGTGTGACAGATGGACCTGGAATTAATAACAACGAGTGCAGTGG 480  
 QY 161 LeuLeuGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180  
 Db 481 CTCAITGAAGGAGCAGCAATAAGATAATAGACITTCGTTTCAATCAATTTTCTCAGAG 540  
 QY 181 CysSerTTPAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200  
 Db 541 TGTAGTGGACCAITTAATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTCT 600  
 QY 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220  
 Db 601 GCATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGACAGTGTCCCTGAGGTGTT 660  
 QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240  
 Db 661 GCCACATCAGGTATGCTTGTGTCATTTTATGATGCTGCTTAAATTTGACTGGA 720  
 QY 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260  
 Db 721 TTTAATATTACTACAGTTTTCATATGTGCCAATNACTGTCTCAGCGCGGAGAGTGT 780

QY 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTTPLysGlyGlu 280  
 Db 781 AAGATCAGTAATAGCAGCGAACTGTTGAATGTGAATGTTCTGAAACTCGAAAGGTGAA 840  
 QY 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyVileCysAsn 300  
 Db 841 GCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTCTCATCGAGGATCTGCAAT 900  
 QY 301 SerSerAspValArgGlyCysSerCysPheSerAspTTPGlnGlyProGlyCysSerVal 320  
 Db 901 TCAAGTGATGTCAGAGGATGCTCTGCTCTCAGACTGGCAGGGTCTCGGATGTTCAGTT 960  
 QY 321 ProValProAlaAsnGlnSerPheTTPThrArgGluGluTyrSerAsnLeuLysLeuPro 340  
 Db 961 CCTGTACAGCTAAACAGTCAATTTGGACTCGAGAGATATTTAACTTTAAAGCTCCCC 1020  
 QY 341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTTPValValGlyGlyTyrMet 360  
 Db 1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTTATGTGGTGTGTGGAGGATATG 1080  
 QY 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTTPLeu 380  
 Db 1081 TTCAACCACTCAGATTATTAACATGGTTCTAGCGTATGACCTTGCTTCTAGGAGTGCTT 1140  
 QY 381 ProLeuAsnArgSerValAsnValValArgTyrGlyHisSerLeuAlaLeuTyr 400  
 Db 1141 CCCTAAACCGCTTCTGTGAACAATGTGGTTGTAGATATGGTCAATCTTTGGCATTATAC 1200  
 QY 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420  
 Db 1201 AAGGATAAAATTTACATGATGGAGGAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
 QY 421 LeuArgValPheHisIleHisAsnGluSerTTPValLeuLeuThrProLysAlaLysGlu 440  
 Db 1261 TTGAGAGTTTTCATCATTAATAGTCATGGGTGTGTGTGACCCCTAAGCGCAAGGAG 1320  
 QY 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460  
 Db 1321 CAGTATGAGTGGTGGGCACTCTGCACACATTTGTTACATGAAGATGGCCGAGTGGTC 1380  
 QY 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480  
 Db 1381 ATGTGTGTCATCTTGTGTCATGCTGCTCTATGGATATATAAGCAATGTGCAGGAATAT 1440  
 QY 481 AspLeuAspLysAsnThrTTPSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500  
 Db 1441 GATTGTGATGAACAACATGGAGTATATTACACCCAGGCGTGCCTTGTGCAAGGGGGT 1500  
 QY 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
 Db 1501 TACGGCCATAGCAGTGTTCAGACCATAGGACCGAGCGGCCCTATACGTTTCATGGTGGCTAC 1560  
 QY 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540  
 Db 1561 AAGCTTTTCAGTGGCAATAGTACCGGCTTCGAGATGATCTTACCGATATGATGTGGAT 1620  
 QY 541 ThrGlnMetTTPThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560  
 Db 1621 ACCCAGATGTGGACCATCTTAAAGCAGCGGATTTTCCGTTACTTTCACACAGCTGTG 1680  
 QY 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
 Db 1681 ATAGTGAGTGGAACTGCTGTTGTTGGGGGAAACACACACAATGACATCATCTATGAGC 1740  
 QY 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTTP 600  
 Db 1741 CATGGCCCAAAATGCTTCTCTCAGATTTTCAGCTTATGACATTTGCTGTGACCGCTGG 1800  
 QY 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620  
 Db 1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTTCAGCAGTC 1860  
 QY 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640

1861 TTACACACAGCACCATGATGTTGTCGGTGGTTTCAATAGTCTCTCTCAGCGACATC 1920  
641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaCysLeuAla 660  
1921 CTGGTATTACCTCGAAGAGTGTGATGCGCATCGGAGTGAAGCGCTTGTATTAGACA 1980  
661 GlyProGlyLeuArgCysValTTPAsnThrGlySerSerGlnCysIleSerTTPAlaLeu 680  
1981 GGAACCTGATTCGGTGTGTTGGACACAGAGTGTCTCAGTGTATCTCGTGGCGCTG 2040  
681 AlaThrAspGluGlnGluGlnLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700  
2041 GCAACTGATGACACAGAGAAAGTTAAATCAGATGTTTTCAAAAGAACTCTTGAC 2100  
701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720  
2101 CATGACAGATGTGACACAGACACAGATTGTTACAGCTGTACAGCCACACCAATGACTGC 2160  
721 HisTTPCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
2161 CACTGGTCAATGACATGTTGTCCTCCAGAGAACACAGCTGTCTCAGAGGCCAGATCTCC 2220  
741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760  
2221 ATTTTAGGTATGAGAAATGCCCCAAGGATAACCCCATGCTACTCTGTACCAAGAGACC 2280  
761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTTPGluProArgAsnGlnCys 780  
2281 AGCTGAGAGCTGTGCTCCAGACAGAACTGCAGTGGAGCCCCCGAATCAGGAGTGC 2340  
781 IleAlaLeuProGluAsnIleCysGlyIleGlyTTPHisLeuValGlyAsnSerCysLeu 800  
2341 ATTGCCCTGCCGAAATATCTGTGCACTGGCTGCAATTTGGTTGGAAACTCATGTTTG 2400  
801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820  
2401 AAAATTACTCTCCAGAGGAATATTGACATGCTAAATGTTCTGTAGGAACCAAT 2460  
821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840  
2461 GCCCTTTGGCTTCTCTTCAACCCAGAGAGAGTAAATTTGTCCTTAAGCAGCTGCA 2520  
841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTTPValGlyLeuArg 860  
2521 ATAATGAGTCATCTCAGAGCATGTCCAGCTCACCTTAACCCATGGCTCGCCCTTCGG 2580  
861 LysIleAsnValSerTyrTTPCysTTPGluAspMetSerProPheThrAsnSerLeuLeu 880  
2581 AAGATCAATGTCTCTACTGCTGGAGATATGTCCTCCATTTACAAATAGTTTACTTA 2640  
881 GlnTTPMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
2641 CAGTGGATGCGCTGTCAGGCCAGTGTGCTGGATTCTGTGGAATTTATCAGAACCCAGT 2700  
901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
2701 ACTCGGGGACTGAAGGCTGCAACTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGG 2760  
921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
2761 CCTGMAAACACAGTCAGTGAAGCAGTGCAGACACCATGTGCTTGGAGCAGCATGTGGA 2820  
941 AspCysThrSerGlySerSerGluCysMetTTPCysSerAsnMetLysGlnCysValAsp 960  
2821 GATTGCAACGAGCGAGCTCTGAGTGCATGTGTGTGAGCAGCAACATGAAGCAGTGTGGAC 2880  
961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTTPTyrThrMetSer 980  
2881 TCCAACTGCTATGTCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2940  
981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000

Db 2941 ACCTGCCCCCTGAAATTTCTCAGGCTACTGTACCTGTAGTCACTGCTTGGAGCAACCA 3000  
Qy 1001 GlyCysGlyTTPCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020  
Db 3001 GGCTGTGGCTGGTGTACTGATCTCCAGCAATATCTGGCAAGGGAATTCATAGAGGCTTC 3060  
Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040  
Db 3061 TATAAAGCAGCAGTGAAGATGCTTCCAGAGCCCTTACAGAGAAATTTCTATCCAGCCC 3120  
Qy 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTTPSerPheIleHisCys 1060  
Db 3121 CTGCTCATTTCAGCATGTCTTAGAGGACAGCAGATACAATCACTGGCTTCTTCACTACTGT 3180  
Qy 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080  
Db 3181 CCAGCTTGCATGCAACCGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAGTGT 3240  
Qy 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100  
Db 3241 GAGAACTGACACAGCAGCAAGCACTGCGAGACTGTGATATCTGGCTTCTACGGTGATCCC 3300  
Qy 1101 ThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120  
Db 3301 ACCAATGGAGGAAATGTGAGCCATGCAATGGCAGCGCTCTCTGTGCAACACC 3360  
Qy 1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140  
Db 3361 AACACGGGCAAGTCTTCTGCACACCAAGGGCGTCAAGGGGACGAGTGCCAGCTATGT 3420  
Qy 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160  
Db 3421 GAGGTAGAAATCGATACCAAGGAACCTCTCAGAGGAACATGTTATTATATCTCTTCT 3480  
Qy 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIle 1180  
Db 3481 ATTGACTATCAGTTTCACTTTAGTCTATCCAGAGATGATCGCTATTACACAGCTATC 3540  
Qy 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200  
Db 3541 AATTTTGGCTACTCTCTGACCAACAAACAGGGATTGGACATGTTCACTCAATGCTCTCC 3600  
Qy 1201 LysAsnPheAsnLeuAsnIleThrTTPAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220  
Db 3601 AAGAAATTTCAACCTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660  
Qy 1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240  
Db 3661 GAAGAGATGCTGTGTGTTTCAAAACCAACATTAAAGAGTACAAAGATAGTTTCTTAAT 3720  
Qy 1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260  
Db 3721 GAGAGTTTGAATTTGCAACACCCCAATATCACTTCTTGTATTATGTAGTAATTTTC 3780  
Qy 1261 ThrTTPProLysIleGlnIleAlaPheSerGln 1272  
Db 3781 ACCTGGCCCAATAAATTCAGGTGCAAACTGAACAA 3816

## RESULT 8

AAZ91921  
ID AAZ91921 standard; cDNA; 4072 BP.

XX AC AAZ91921;

XX DT 08-JUN-2000 (first entry)

XX DE Human mahogany protein coding sequence #3.

XX KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
KW weight regulation; cell therapy; body weight disorder; cachexia;  
KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.

OS	Homo sapiens.	Db	226	CTGCCCCGGAGGAGCGCGCGCGTGGTGGCGCGCGCGTGGCGCGCGCGCGC	285
XX					
PN	WO200005373-A2.	Qy	97	GlualalysGluCysaspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly	116
XX					
PD	03-FEB-2000.	Db	286	GAGGCCAAGGAATGTGACCGCGCGTGTGTCAACGGCGGTGCTGCAACCCCTGGCACCGGC	345
XX					
PF	21-JUL-1999; 99WO-US016484.	Qy	117	GluCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe	136
XX					
PR	21-JUL-1998; 98US-0093630P.	Db	346	CAGTGGCTGTGCCCGCGCGTGGTGGCGGCAATGCCAGCACTGCGGGGCGCGCTTC	405
PR	20-OCT-1998; 98US-0104978P.				
PR	05-FEB-1999; 99US-00245041.	Qy	137	ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTrpLysThr	156
XX					
PA	(MILL-) MILLENIUM PHARM INC.	Db	406	AGACTAACTGGATCTCTGGGTTTGTGACAGATGGACCTGGAATATATAATACRAAACG	465
XX					
PI	Moore K, Nagle DL;	Qy	157	LysCysThrTrpLeuLeuGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis	176
XX					
DR	WPI; 2000-195103/17.	Db	466	AAGTGACAGTGGCTCATTTGAAGGACAGCCAAATAGAATAATAGAGACTTCGTTTCAATCAT	525
DR	P-PSDB; AAY81808.				
XX	New human and murine mahogany genes, useful, e.g. for diagnosis and	Qy	177	PheAlaThrGluCysSerTrpAspHisLeuTyValTyAspGlyAspSerIleTyAla	196
PT	treatment of body weight disorders.				
XX		Db	526	TTTGCTACAGATGTAGTTGGACCATTTATATGTTTATGATGGGACTCAATTTATGCA	585
XX	Claim 1; Fig 19a; 188pp; English.				
PS	This sequence represents a human mahogany gene of the invention. The	Qy	197	ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal	216
CC	mahogany genes are used: (i) to produce recombinant mahogany (mg)				
CC	proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming	Db	586	CCGCTAGTCTGCATTTAGTGGCTCATTTGCTGAGAGAGATGGCAATGAGACTGTC	645
CC	therapeutics; (iii) as a source of diagnostic probes and primers for	Qy	217	ProGluValValAlaThrSerGlyTyAlaLeuLeuHisPhePheSerAspAlaAlaTy	236
CC	detecting expression of mg genes or mutations, regulatory defects, in				
CC	this gene, or for isolation of related sequences; and (iv) in (cell-	Db	646	CTGAGGTGTTGCCACATCAGGTTATGCTTCTGCTCATTTTITTAGTGATGCTGCTTAT	705
CC	based) gene therapy. (ii) are used to raise specific antibodies (Ab); to	Qy	237	AsnLeuThrGlyPheAsnIleThrTySerPheAspMetCysProAsnAsnCysSerGly	256
CC	identify other (extra)cellular products involved in weight regulation,				
CC	and to screen for agents that disrupt interaction between (ii) and other	Db	706	AAATTGACTGGATTAAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	765
CC	macromolecules. The Ab are used to detect abnormal levels (or function)	Qy	257	ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn	276
CC	of (ii) for diagnosis, prognosis or monitoring of treatment); to				
CC	evaluate (iii)-expressing cells intended for cell therapy, and as	Db	766	CGAGGAGAGTGAAGATCAGTAATAGCAGCGATCTGTTGAATGTGAATGTGTTCTGAAAC	825
CC	therapeutic mg inhibitors. Cells that express the mg gene (or contain the	Qy	277	TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg	296
CC	mg polypeptide) are used to identify agents (A) that modulate mg				
CC	activity. (A) are potentially useful for the treatment of body weight	Db	826	TGGAAGAAGTGAGCATGTGACATCTCTACTGTACAGACAACTGTGGTTTCTCTCATCGA	885
CC	disorders, particularly obesity, cachexia or anorexia, or other	Qy	297	GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro	316
CC	conditions associated with the mg gene such as hyperpigmentation,				
CC	hyperphagia and disorders that result in increased metabolic rate	Db	886	GGCATCTGCAATTCAAGTGATGTACAGAGGATGCTCTGCTTCTCAGACTGGCAGGGTCT	945
XX	Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;				
SQ					
Alignment Scores:					
Pred. No.:	0	Length:	4072		
Score:	6870.50	Matches:	1240		
Percent Similarity:	97.41%	Conservative:	3		
Best Local Similarity:	97.18%	Mismatches:	20		
Query Match:	86.80%	Indels:	13		
DB:	3	Gaps:	3		
US-09-787-097-12 (1-1429) x AAZ91921 (1-4072)					
Qy	2	ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAlaAla	21		
Db	10	GTGGCCGAGCGCGCGCAACTGAGCAAGGCTGAGAGAGAGCGCGCGCGCGCGCGC	69		
Qy	22	LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro	41		
Db	70	CTCGCGGCGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	129		
Qy	42	GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro	56		
Db	130	GGACCGCGCGCGCGGTGTGTCTCCCGCGGGTGTCTCGCGGCGCTGCTCGCGGCGCG	189		
Qy	57	LeuArgProArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu	76		
Db	190	CTGCTGCGG---CTGCTCTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCG	225		
Qy	77	LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAla	96		

QY 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476  
DB 1366 GCGCCAGTGGTCACTGGTCACTTTGGTCACTGCCCTCTATGGATATATAGCAAT 1425  
QY 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496  
DB 1426 GTGCAGGAATATGATTTGGATAAGAACACATGGAGTATATATACACACCGGGTGGCCCTT 1485  
QY 497 ValGlnGlyGlyTyrGlyHisSerValTyrAspHisArgThrArgAlaLeuTyrVal 516  
DB 1486 GTGCAGGGGGTTACGCCATAGCAGTGTTCACGACCATAGGACACGAGGCCCTTATACGTT 1545  
QY 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspLeuTyrArg 536  
DB 1546 CATGGTGGCTACAGGCTTTCACTGCCAATAAGTACCGGCTTGACAGATGATCTCTACCGA 1605  
QY 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPheArgTyrLeu 556  
DB 1606 TATGATGTGGATACCCAGATGGACCATCTTTAAGGACAGCCGATTTTCCGTTACTTG 1665  
QY 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576  
DB 1666 CACACAGCTGTATAGTGAGTGAACCATGCTGTGGTGTGGGGAAACACACACATGAC 1725  
QY 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596  
DB 1726 ACATCTATGAGCGCATGGCGCAAAATGCTTCTTCAGATTTTCAGGCTATGACATTGCC 1785  
QY 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616  
DB 1786 TGTGACCGCTGTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAAAGATTGGC 1845  
QY 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
DB 1846 CATTCAGCAGTCTTACACACAGCACCATGATGTGTTCGGTGGTTTCAATAGTCTCCTC 1905  
QY 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
DB 1906 CTCAGCGCATCCTGTGATTCACCTCGGAACAGTGTGATGGCATCGAGTGAGCCGCT 1965  
QY 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676  
DB 1966 TGTTTAGCAGCAGGACCTGGTATTTCGGTGTGTGGAAACACAGAGTCTGCTCAGTGTATC 2025  
QY 677 SerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696  
DB 2026 TCGTGGCGCTGGCACTGATGAACAGAGAAAAGTTAAATCAGATGTTTTTCCAAA 2085  
QY 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
DB 2086 AGNACTTTGACCATGACAGATGTGACCCAGCACACAGATTTTACAGCTGCACAGCCAA 2145  
QY 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
DB 2146 ACCAATGACTGCCACTGGTCAATCACCATTGTGTCGCCAGGAACACACAGTGTCTCAGAA 2205  
QY 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
DB 2206 GGCAGATCTCCATTTTATAGTATGAGATTTGCCCCAGGATAACCTATGTACTACTGT 2265  
QY 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776  
DB 2266 AACAGAAGACAGCTGCAGGAGCTGTGCCCTGGACAGAACTGCAGTGGAGGCCCGG 2325  
QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796  
DB 2326 AATCAGAGTGTGATTCGCTGCCGGAATAATCTGTGGCATTTGGCTGGCATTTGGTTGA 2385  
QY 797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
DB 2386 AACTCATGTTTGAATAATTACTACTGCCAAGGAGAAATTATGACAAATGCTAAATTTGTTCT 2445

QY 817 ArgAsnHisAsnAlaLeuLeuAspSerLeuThrThrGlnLysLysValGluPheValLeu 836  
DB 2446 AGGAACACCAATGCCCTTTTGGCTTCTCTTACAACCCAGAAAGAGTAGAATTTGTCTCT 2505  
QY 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856  
DB 2506 AAGCAGCTCGAATAATGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGG 2565  
QY 857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876  
DB 2566 GTCCGCTTCGGAAGATCAATGTGTCTACTGCTGGGAGAGATATGTCCCATTTTACA 2625  
QY 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
DB 2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCAGCATGCTGGATTCTGTGGAAATTTTA 2685  
QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
DB 2686 TCAGAACCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745  
QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
DB 2746 GTCTGTGAAGGCTGCAACCAACAGTGTAAAGAGTCCGGACACCATGTCCCTTGAGG 2805  
QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956  
DB 2806 ACAGCATGTGGAGTTGACCCAGCGGACGCTCTGAGTGCATGTGTGGTGCAGCAACATGAAG 2865  
QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrp 976  
DB 2866 CAGTGTGTGACTCCCAATGCCTATGTGGGCTCTCTCCCTTTTGGCCAGTGTATGAATGG 2925  
QY 977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
DB 2926 TATACGATGAGCACTTGGCCCCCTGAAAATTTGTTCAGGCTACTGTACTGTAGTCATTGC 2985  
QY 997 LeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016  
DB 2986 TTGGAGCAACAGGCTGTGGTGTACTGTATCCAGCAATACTGGCAAGGAAAGAAATGC 3045  
QY 1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036  
DB 3046 ATAGAGGGTTCCTATAAAGACACAGTGAAGATGCTTCCAAAGCCCTCAGAGAAATTC 3105  
QY 1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056  
DB 3106 TATCCACAGCCCTGTCTCAATTCACCATGTGTCTAGAGGACAGCAGATACAACTGGTCT 3165  
QY 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
DB 3166 TTCAATCACTGTCCAGCTTGGCAATCAACGGCCACAGTAAATGCATCAATCAGAGCATC 3225  
QY 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096  
DB 3226 TGTGAGAGTGTGAGAACTGACCA CAGGCAAGCATCTGCGAGACCTGCATATCTGGCTTC 3285  
QY 1097 TyrGlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116  
DB 3286 TAGCGTGTATCCCAATCGGAGGAAATGTCAAGCCATGCAAGTGCATGGCAGCGCTCT 3345  
QY 1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136  
DB 3346 CTGTGCAACACCAACACGGGCAAGTCTCTGACCCACCAAGGGCGCTCAAGGGGACGAG 3405  
QY 1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156  
DB 3406 TGCCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCCCTCTCAGAGAAACATGTTAT 3465  
QY 1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176  
DB 3466 TATACTCTTCTTATGACTATCAGTTACCTTTAGTCTATCCACGAGAAAGATGATCCCTAT 3525  
QY 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196

```

Db 3526 TACACAGCTATCAATTTTGTGCTACTCTGACGAAACAAACAGGATTCGACATGTC 3585
QY 1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly 1216
Db 3586 ATCANTGCTCCAGAAATTCACCTCAACATCAGCTGGCTGCCAGTTCTCAGCTGA 3645
QY 1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp 1236
Db 3646 ACCCAGGCTGGAGAGAGATGCTCTGTTTCAAAACCAACATTAAGGAGTCAAAAGAT 3705
QY 1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256
Db 3706 AGTTTCTCTAAGAGAAGTTGATTTTTCGACACACCCCAATATACATTTCTTTGTTAT 3765
QY 1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
Db 3766 GTCAAGTAATTCACCTGGCCCATCAAAATTCAGGTGCAACAACTGAACAA 3813

RESULT 9
AAZ52274
ID AAZ52274 standard; cDNA; 3597 BP.
XX AC AAZ52274;
DT 18-JUL-2000 (first entry)
XX DE Human soluble attractin-1 cDNA.
XX KW Human; soluble attractin-1; immune response; macrophage; monocyte;
KW T cell; immunostimulant; immunosuppressed patient; cancer;
KW immunodeficiency syndrome; transplant; autoimmune disease; ss.
XX OS Homo sapiens.
XX FH Key
XX CDS 1..3597
XX FT Location/Qualifiers
XX FT 1..3597
XX FT /*tag= a
XX FT /product= "Soluble attractin-1"
XX PN WO200015651-A1.
XX PD 23-MAR-2000.
XX PF 14-SEP-1999; 99WO-US020948.
XX PR 14-SEP-1998; 98US-0100137P.
XX PA (DAND ) DANA FARBER CANCER INST INC.
XX PI Duke-Cohan JS, Schlossman SF;
XX DR WPI; 2000-271373/23.
XX DR P-PSDB; AA70689.
XX PT Isolated nucleic acids encoding human attractin polypeptides useful for
XX FT enhancing immune responses.
XX PS Claim 3; Fig 8; 120pp; English.
XX CC The patent discloses four forms of human attractin polypeptides which
XX CC enhance immune response by promoting macrophage and monocyte spreading in
XX CC the presence of T cells. These include soluble attractin-1 and -2 and
XX CC membrane attractin-1 and -2. These various forms of attractin are encoded
XX CC by alternatively spliced mRNA molecule transcribed from a single gene.
XX CC The present sequence is a cDNA encoding soluble attractin-1 obtained from
XX CC activated T cell and human foetal liver libraries. This sequence can be
XX CC used to enhance immune responses in immunosuppressed patients such as
XX CC those undergoing chemo- and radio-therapy treatment for cancer or those
XX CC suffering from common variable immunodeficiency syndrome. The proteins
XX CC may also be used to screen modulators (agonists and antagonists) of
XX CC immune responses which may also be used to regulate immune reactions.
XX CC Attractin antibodies can be used to inhibit immune response in transplant

```

```

CC recipients or patients afflicted with autoimmune disease
XX SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 0 Length: 3597
Score: 6612.00 Matches: 1191
Percent Similarity: 93.72% Conservative: 2
Best Local Similarity: 93.56% Mismatches: 5
Query Match: 83.54% Indels: 76
DB: 3 Gaps: 1
US-09-787-097-12 (1-1429) x AAZ52274 (1-3597)
QY 1 MetValAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAla 20
Db 1 ATGGTGGCGGCGGCGGCACTGAGGCAAGGCTGAGGAGGAGGCGGCGGCGGCA 60
QY 21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTyrAspValThrArgAlaGlyArg 40
Db 61 GCGCTGCGGGCAGGAGCGGCGGCGGCGC----- 90
QY 41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
Db 90 ----- 90
QY 61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80
Db 90 ----- 90
QY 81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100
Db 90 ----- 90
QY 101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
Db 91 -----TGTCACACGCGGCTGCTGCAACCTGGCAGCCCGGCGGCGGCTGC 138
QY 121 ProAlaGlyTrpValGlyGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140
Db 139 CCGCGCGGCTGGTGGCGGCGGCAATGCCAGCACTGCGGGGCGGCTTCAGACTAATCGA 198
QY 141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160
Db 199 TCTTCTGGGTTTGTGACAGATGACCTGGAAATATATAATACAAAACGAAGTGCAGCTGG 258
QY 161 LeuIleGluGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
Db 259 CTCATTGAAGGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATTTTGTCTACAGAG 318
QY 181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200
Db 319 TGTAGTTGGACCATTTATATGTTTATGATGGGACTCAATTTATCACCCTAGTTGCT 378
QY 201 AlaPheSerGlyLeuLeuValProGluArgAspGlyAsnGluThrValProGluValVal 220
Db 379 GCATTAGTGCGCTCAATGTTCTGAGAGATGGCAATGAGACTTCCTCCCTGAGTTGTT 438
QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240
Db 439 GCCACATCAGTTATGCTTGTGCAATTTTGTGATGCTGCTTATATTTGACTGGA 498
QY 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
Db 499 TTTAAATATTACTTACAGTTTGTATATGTCTCCAAATAACTGCTCAGGCGGAGAGAGTGT 558
QY 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280
Db 559 AAGATCAGTAATAGCAGCGCATCTGTTGAATGTGAATGTTCTGAAACCTGGAAAGGTGAA 618
QY 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyLysCysAsn 300
Db 619 GCATGTGACATTCCTCACTGTACAGACAACTGTGGTGTTCCTTCATCGAGGCACTGCAAT 678

```



301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320  
Db TCAAGTGATGTCAGAGGATGCTCTGCTCTCAGACTGGCAGGGTCTGGAGTTTCAGTT 738  
321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLeuPro 340  
Db CTGTGTACAGCTAACAGCTCATTTTGGACTCGAGAGGAATATTCTTAACCTTAAGCTCCCC 798  
341 ArgAlaSerHisLeuAlaValAlaValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360  
Db AGAGCATCTCTAATAAGCTGTGCTCAATGGAACAATATTATGTGGTGTGTTGGAGGATATATG 858  
361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380  
Db TTCAACCACTCAGATATAACATGCTTCTAGGGTATGACCTTGTCTTAGGAGGTGGCTT 918  
381 ProLeuAsnArgSerValAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400  
Db CCACTAAAACCGTCTCTGTGAACAATCTGCTTGTAGATATGCTCATTTCTTGGCATTATAC 978  
401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420  
Db AAGGATAAAATTTACATGATATGGAGAAATTTGATTCACTGGGAATGTACCAATGAG 1038  
421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440  
Db TTGAGAGTTTTTCACATTCATTAATGAGTCATGGGTGTGTTGTCACCCCTTAAGCAAGAGG 1098  
441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460  
Db CAGTATGCACTGGTGGGCACCTCTGCACACATTTGTTACCAAGAATGGCGGAGTGGTC 1158  
461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnLeuTyr 480  
Db ATGCTGGTCATCTTGGTCACCTGCCCTCTATGATATATAAGCAATGTGCAGGAATAT 1218  
481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500  
Db GATTTGGATAAAGACACATGGAGTATATTACACACCCAGGGTGGCCCTTGTCAAGGGGGT 1278  
501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
Db TACGGCCATAGCAGTGTGTTACGACCATAGGACCGGCCCTTATAGCTTCATGGTGGCTAC 1338  
521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540  
Db AAGGCTTTTCAGTGCCTAATTAAGTACCGGCTTGACATGATCTCTACCGATATGATGTGGAT 1398  
541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560  
Db ACCAGATGTGACCACTCTTAAGGACAGCCGATTTTCCGGTTACTTGCACACAGCTGTG 1458  
561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
Db ATAGTGAGTGAACCACTGCTGTGTTGGGGGAAACACACACAATGACACATCTATGAC 1518  
581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600  
Db CATGGCGCAAAATGCTCTCTTCATGATTTACGGCTATGACATGTCCTGTGACCGCTGG 1578  
601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaVal 620  
Db TCAGTGTCTCCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCCAGCAGT 1637  
620 IleuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640  
Db CTTACACAACAGCAGCACCATGATGTGTGCTGGTGGTTCATATAGTCTCTCTCCAGCGACAT 1697  
640 IleuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAl 660  
Db CCTGGTATTTCACCTCGGAACAGTGTGATCGGCATCGGAGTGAAGCGCTTGTATTAGCAGC 1757

660 aglyProGlyIleArgCysValTrpAsnThrGlySerGlnCysIleSerTrpAlaLeu 680  
Db AGGACCTGGTATTTCGGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCTGGGCGCT 1817  
680 ualaThrAspGluGlnGluGlyLysLeuLysSerGluCysPheSerLysArgThrLeuAs 700  
Db GGCACTGATGAACAAGAAAGTAAATCAGAATGTTTTTCCAAAGAACTCTTGA 1877  
700 PHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspC 720  
Db CCATGACATGTGACACGACACACAGATTGTACAGTGCACAGCCACCAACCATGACTG 1937  
720 SHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSe 740  
Db CCACCTGGTCAATGACCATTTGTCTCCAGAGAACACACAGCTCTCAGAAGGCCACATCTC 1997  
740 rIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysIleS 760  
Db CATTTTAGTATGAGAAATGGCCCAAGGATAACCTATGTACTACTGTAAACAAGAC 2057  
760 rSerCysA-gSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCy 780  
Db CAGCTGCAGAGCTGTGCCCTGGACCCAGAACTGCCAGTGGAGCCCGGAATCAGAGTG 2117  
780 sIleAlaLeuProGluAsnIleCysGlyTleGlyTrpHisLeuValGlyAsnSerCysLe 800  
Db CATTTGCCCTGGCCGAAATATCTGTGGCAATGGCTGGCAATTTGGTTGGAAACTCATGTTT 2177  
800 uLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAs 820  
Db GAAATTTACTACTGCCAGAGGAATTTATGCAATGCTAAATTTGTTCTGTAGGAACCAAA 2237  
820 nAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuAr 840  
Db TGCCTCTTTGGCTCTCTTACAAACCCAGAAAGGAGTAGAATTTGCTTAAAGCAGCTGG 2297  
840 gIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuAr 860  
Db AATTAATGCACTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGCCTTCG 2357  
860 gLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLe 880  
Db GAAGATCAATGTCTCTACTGTGTCTGGGAAGATATATGCCCATTTTACAAATAGTTTACT 2417  
880 uGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSe 900  
Db ACAGTGGATGCCCTCTGAGCCCGATGATGCTGATTTCTGTGGAATTTTATCAGAACCCAG 2477  
900 rThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluAr 920  
Db TACTCGGGGACTCAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAG 2537  
920 gProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGl 940  
Db GCCTGCAAAACACAGTGTCTAAGCAGTGGCCGACACCATGTGCTTGAGGACAGCATGTGG 2597  
940 YAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAs 960  
Db AGATTGCACACAGCGGCGAGCTCTGAGTGCATGTGTGCAGCAACATGAAGCAGTGTGTGA 2657  
960 pSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSe 980  
Db CTCCAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGTATACGATGAG 2717  
980 rThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPr 1000  
Db CACTGCCCCCTGAAATTTGTTCAAGTACTGTACTCTGTAGTCTTGTCTGGAGCAACC 2777  
1000 oGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluLys 1020  
Db AGGCTGTGGCTGTACTGATCCCAAGCAATCTGCAAGGGAATGTCATAGAGGGTTTC 2837  
1020 rTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPr 1040



```
Db 2838 CTATAAGGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCC 2897
QY 1040 oLeuLeuAsnSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCy 1060
Db 2898 CTTGCTCAATTCAGCATGTGCTAGAGGACAGATACACTGCTCTTTCATCTACTG 2957
QY 1060 sProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCy 1080
Db 2958 TCCAGCTTGGCAATGCAAGGCGCACAGTAAATGCATCAATCAGAGCATCTGTGAGAACTG 3017
QY 1080 sGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPr 1100
Db 3018 TGAGAACCTGACCAAGGCAAGCACTGCGAGACCTGTCATATCTGCTCTTCTACGGTGATCC 3077
QY 1100 oThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnTh 1120
Db 3078 CACCAATGAGGAAATGTCAGCCATGCAAGTGCAATGGCAGCGGTCTCTGTGCAACAC 3137
QY 1120 rAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCy 1140
Db 3138 CAACACGGGCAAGTCTTCTGACCAACCAAGGCGCTCAAGGGGGACGAGTCCAGCTATG 3197
QY 1140 sGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLe 1160
Db 3198 TGAGGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTTCT 3257
QY 1160 uIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180
Db 3258 TATTACTATCATGTTTACCTTTAGTCTATCCAGGAGATGATGCTATTACAGCTAT 3317
QY 1180 eAsnPheValAlaThrProAspGluGluAsnArgAspLeuAspMetPheIleAsnAlaSe 1200
Db 3318 CAATTTTGTGGTACTCTCTGACCAACAAACAGGGATTTGGACATGTTTCATCATGCTC 3377
QY 1200 rLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGl 1220
Db 3378 CAAGAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTG 3437
QY 1220 yGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAs 1240
Db 3438 AGAAGAGATGCTGTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAA 3497
QY 1240 nGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPh 1260
Db 3498 TGAGAAGTTTGATTTTCGCAACCAACCAATATATCATCTTCTTTGTTTATGTCAGTAAT 3557
QY 1260 eThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
Db 3558 CACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3594
RESULT 10
AAS72657
ID AAS72657 standard; cDNA; 3597 BP.
XX
AC AAS72657;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #8461.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
```

```
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
XX
DR P-PSDB; ABG08470.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 8461; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: the sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;
```

```
Alignment Scores:
Pred. No.: 0 Length: 3597
Score: 6612.00 Matches: 1191
Percent Similarity: 93.72% Conservative: 2
Best Local Similarity: 93.56% Mismatches: 5
Query Match: 83.54% Indels: 76
DB: 5 Caps: 1
```

US-09-787-097-12 (1-1429) x AAS72657 (1-3597)

```
QY 1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAla 20
Db 1 ATGTTGGCCGACGCGCGCACTGAGGCNAGGCTGAGGAGGAGGACGCGGCGCGGCA 60
QY 21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpValThrArgAlaGlyArg 40
Db 61 GCGCTCGCGGACGAGGCGGCGGCGGCAC----- 90
QY 41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
Db 90 ----- 90
QY 61 LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGlu 80
Db 90 ----- 90
QY 81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaGluAlaLysGlu 100
Db 90 ----- 90
QY 101 CysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys 120
Db 91 -----TGTGTCAACGCGGTGCTGTCAACCTTGGCACCAGCCAGTGGCTCTGC 138
```

QY 121 ProAlaGlyTrrpValGlyGluClnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140  
Db 139 CCCGCCGGCTGGGTGGGGAGCAATGGCCAGCACTCGGGGGCCGCTTCAGACTAACTGGA 198  
QY 141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrp 160  
Db 199 TCTTCTGGGTGTGCACAGATGCACCTGGAAATTAATAACAAAACGAAGTCACAGTGG 258  
QY 161 LeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180  
Db 259 CTCATTGAAGGACGCCCAATAAGATAGACTTCGTTTCAATCATTTTGGCTACAGAG 318  
QY 181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200  
Db 319 TGTAGTTGGGACCATTTATATGTTTATGATGGGACTCAATTTATGACACCGCTAGTTGCT 378  
QY 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220  
Db 379 GCATTTAGTGGCTCATTTGTTTCCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT 438  
QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240  
Db 439 GCCACATCAGGTATGCTTGTGCTATTTTATGATGCTGCTTATTAATTTGACTGGA 498  
QY 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyCys 260  
Db 499 TTTAATATTACTACAGTTTTCATATGTTGTCACAAATAACTGCTCAGGCCGAGGAGTGT 558  
QY 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyCln 280  
Db 559 AAGATCAGTAATAGCAGCGCACTGTTGAATGTGAATGTTCTGAAACTGGAAGGTGAA 618  
QY 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300  
Db 619 GCATGTGACATTCCTCACTGTACAGCAACTGTGGTTTTCCTCATCGAGGCATCTGCAAT 678  
QY 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320  
Db 679 TCAAGTGATGTACAGAGATGCTCCTGCTTCTCAGACTGCGCAGGGTCTCGGATGTTCAAGTT 738  
QY 321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340  
Db 739 CTTGTACCACTCAACCGTCATTTTGGACTCGAGAGGAATATCTTAATTAAGCTCCCC 798  
QY 341 ArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValValGlyClyTyrMet 360  
Db 799 AGAGCATCTCATAAAGCTGTGGTCAATGGAACATATTATGTGGGTTGTTGGAGGATATATG 858  
QY 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380  
Db 859 TTCACCACTCAGATTATACATGTTCTAGCGATGATGACCTTGCTTCTTAGGGAGTGGCTT 918  
QY 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400  
Db 919 CCACATAACCGTTCTGTGAACAATGTGGTTGTGTAGATATGTTCAATCTTGGCATATATAC 978  
QY 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420  
Db 979 AAGGATAAATTTACATGATATGAGGAAAAATGTAATCAACTGGGAATGTGACCAATGAG 1038  
QY 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440  
Db 1039 TTGAGAGTTTTCACATTCATTAATGATCATGGTGTGTTGTTGACCCCTAAGGCAAGGAG 1098  
QY 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460  
Db 1099 CAGTATGAGTGGTGGGCACTCTGCACACATTTGTTACATGAAGAATGGCCGAGTGTCT 1158  
QY 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480  
Db 1159 ATGCTGGTTCATCTTTGGTCACTGCCCTCTCTATGGATATATAGCAATGTGCGAGGATAT 1218

QY 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500  
Db 1219 GATTTCGATTAAGAACACATAGGAGTATATTACACACCCAGGGTGCCTTGTGCAAGGGGT 1278  
QY 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
Db 1279 TACGGCCATAGCAGTGTTCACGACCATAGACAGGGCCCTATACGTTTCATGGTGCTAC 1338  
QY 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540  
Db 1339 AAGGCTTTTCAGTCCCAATAGTACCGCTTGCAGATGATCTCTACGATATGATGTGGAT 1398  
QY 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560  
Db 1399 ACCCAGATGTGACCATTTCTTAAGGACAGCCGATTTTTCGTTACTTGCACACAGCTGG 1458  
QY 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
Db 1459 ATAGTGAGTGGAAACCAATGCTGCTGTTTGGGGGAAACACACACAATGACACATCTATGAC 1518  
QY 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600  
Db 1519 CATGGCCCAAAATGCTTCTTTCAGATTTTCATGGCCCTATGACATTCCTGTGACCGCTGG 1578  
QY 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620  
Db 1579 TCAGTGTCTTCCCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTCACAGCT 1637  
QY 620 IleHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640  
Db 1638 CTTTACACACAGCACCATGATGTGTTGCTGTTTCAATAGTCTCCTCCTCAGCGACAT 1697  
QY 640 eLeuValPheThrSerGluClnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaIle 660  
Db 1698 CTTGTTATTCACCTCGAAACAGTGTGATCGCATCGGAGTGAAGCCGCTGTTTACGACG 1757  
QY 660 aglyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLe 680  
Db 1758 AGGACCTGTTATTCGTTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGCGCT 1817  
QY 680 uAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLysArgThrLeuAs 700  
Db 1818 GGCAACTGTATGAACAAGAAAAAGTAAAAATCAGAAATGTTTTTCCAAAAGAACTCTTCA 1877  
QY 700 pHHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspC 720  
Db 1878 CCATGACAGATGTGACGACACACAGATTTTACAGTGTGACAGCCACACCACTAGCTG 1937  
QY 720 sHIStrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyClnIleSe 740  
Db 1938 CCAGTGTGCAATGACCAVTTGTGTCGCCAGGAACACACAGCTGCTCAGAGAGCCAGATCTC 1997  
QY 740 rIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysTh 760  
Db 1998 CATTTTAGTATGAGAAATTTGCCCAAGGATAAACCTATGTACTACTGTAAACAGAGAC 2057  
QY 760 rSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluC 780  
Db 2058 CAGCTGCAGGAGCTGTGCTCGACAGAACTGCCAGTGGAGCCCGGAAATCAGGAGTG 2117  
QY 780 sIleAlaLeuProGluAsnIleCysGlyIleGlyTyrHisLeuValGlyAsnSerCysLe 800  
Db 2118 CATTTGCCCTCCCGAAAAATATCTGTGCGATTTGCTGGCATTTTGGTGGAAACTCATGTTT 2177  
QY 800 uLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAs 820  
Db 2178 GAAATTTACTTCTGCCCAAGGAGAAATATGACATGTAAATTTCTTCTGTAGGAACCAACA 2237  
QY 820 nAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuAr 840  
Db 2238 TGCCCTTTTGGCTTCTCTACAAACCCAGAGAGGAGGATTTGCTTAAAGCAGCTGGC 2297  
QY 840 gIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuAr 860

Db 2298 AATAATGCGATCATCTCAGAGCATGTCACCTTAACCCCATGGGTGGCGCTTGG 2357  
QY 860 gylsileAenValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLe 880  
Db 2358 GAAGATCAATGTGTCTACTGTGTGCTGGAGATATGTCCCATTTACAAATAGTTTACT 2417  
QY 880 uGlnTrpMetProSerGluProSerAspIadlyPheCysGlyIleLeuSerGluProSe 900  
Db 2418 ACAGTGGATGCCCTCTGAGCCCATGTATGCTGATTTCTGTGAAATTTATACAGACCCAG 2477  
QY 900 rThrArgGlyLeuLeuAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluAr 920  
Db 2478 TACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGTAGTGTCTGTGAAG 2537  
QY 920 gProAlaAenHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGI 940  
Db 2538 GCTGTCAAAACCAAGTCTTAAGCAGTGCAGCGGACACCATGTGCTTGCAGACAGCATGTGG 2597  
QY 940 yAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValas 960  
Db 2598 AGATTGCACAGCGGCAGCTCTGAGTGCATGTGTGCAGCAACATGAAGCAGTGTGTGA 2657  
QY 960 pSerAenAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSe 980  
Db 2658 CTCCAATGCTATGTGGCTCTCTCCCTTTGCGCAGTGTATGTAATGGTATACCATGAG 2717  
QY 980 rThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluLapR 1000  
Db 2718 CACCTGCCCTCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACC 2777  
QY 1000 oGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySe 1020  
Db 2778 AGGCTGTGGCTGGTGTACTGATCCAGCAATACTGGCAAGGGAATAATCATAGAGGGTTC 2837  
QY 1020 rTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPr 1040  
Db 2838 CTATAAGGACCAAGTGAAGATGCCCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCC 2897  
QY 1040 oLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCy 1060  
Db 2898 CCTGTCTCAATCCAGCATGTCTAGAGGACAGAGATACACTGGTCTTTCATTCACGTG 2957  
QY 1060 sProAlaCysGlnCysAenGlyHisSerLysCysIleAenGlnSerIleCysGluLysCy 1080  
Db 2958 TCCAGCTTCCCAATCAACGGCCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTG 3017  
QY 1080 sGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPr 1100  
Db 3018 TGAGAACTGACCAAGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGATCC 3077  
QY 1100 cThrAenGlyLysCysGlnProCysLysCysAenGlyHisAlaSerLeuCysAenTh 1120  
Db 3078 CACCAATGGAGGAAATGTACGCAATGCAAGTGCATGGCAAGGCTCTCTGTGCACAC 3137  
QY 1120 rAsnThrGlyLysCysPheCysThrTrpLysGlyValLysGlyAspGluCysGlnLeuCy 1140  
Db 3138 CAACACGGGCAAGTCTCTGACCAACCAAGGGGTCAAGGGGACGAGTGCACAGCTATG 3197  
QY 1140 sGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTrpThrLeuLe 1160  
Db 3198 TGAGTAGAATAATCGATACCAAGGAACCTCTCAGAGACATGTTATATATCTTCT 3257  
QY 1160 uIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTrpThrAlaIl 1180  
Db 3258 TATTACATCATGTTTCACTTTAGTCTATCCAGAGAGATGATCGCTATTACACAGCTAT 3317  
QY 1180 eAsnPheValAlaThrProAspGluInAsnArgAspLeuAspMetPheIleAenAlaSe 1200  
Db 3318 CAATTGTGCTACTCTCTGACGACAAACAGGATTTGGACATGTTTCATCATGCTC 3377  
QY 1200 rLyAsnPheAsnLeuAenIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGI 1220

Db 3378 CAAGAATTTCAACCTCAACATCACTGGTGGTGGCAGTTTCTCAGCTGAACCCAGGCTGG 3437  
QY 1220 yGluGluMetProValValSerLysThrAsnIleLysGluTrpLysAspSerPheSerAs 1240  
Db 3438 AGAAGAGATGCTGTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAA 3497  
QY 1240 nGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPh 1260  
Db 3498 TGAGAAGTTTGATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTT 3557  
QY 1260 eThrTrpProLysLysIleGlnIleAlaPheSerGln 1272  
Db 3558 CACCTGCCCCATCAAAATTCAGGTGCACAACTGAACAA 3594  
RESULT 11  
ADD70997  
ID ADD70997 standard; DNA; 3597 BP.  
XX AC ADD70997;  
XX 15-JAN-2004 (first entry)  
XX Human attractin gene SEQ ID NO:1.  
DE liver cancer; chronic hepatitis; cirrhosis; liver disease; hepatotropic;  
XX cytotstatic; gene therapy; human; gene; ds.  
KW Homo sapiens.  
XX WO2003061564-A2.  
XX 31-JUL-2003.  
XX 20-DEC-2002; 2002WO-US040718.  
XX 21-DEC-2001; 2001US-0341815P.  
XX 31-DEC-2001; 2001US-0343185P.  
XX (GENE-) GENE LOGIC INC.  
XX (LGBI-) LG BIOMEDICAL INST.  
XX Koh SS, Liu Q, Chung H, Zeng W, Lee B, Yeramilli S, Song SY;  
XX WPI; 2003-663343/62.  
XX Diagnosing liver cancer cells, useful for treating liver cancer  
XX associated with chronic hepatitis or cirrhosis comprises detecting the  
XX level of expression in a tissue sample of one or more genes associated  
XX with cancerous liver tissues.  
PS Claim 1; SEQ ID NO 1; 176pp; English.  
XX The present invention describes a method for diagnosing liver cancer  
XX cells comprising detecting the level of expression in a tissue sample of  
XX one or more genes given in the specification (see ADD70997 to ADD71105),  
XX where differential expression of the genes is indicative of liver cancer.  
XX Also described: (1) detecting the progression of liver cancer in a  
XX patient; (2) monitoring the treatment of a patient with liver cancer; (3)  
XX treating a patient with liver cancer; (4) typing a liver disease in a  
XX patient; (5) detecting the presence or progression of liver cancer in a  
XX patient with chronic hepatitis or cirrhosis; (6) differentiating liver  
XX cancer related to chronic hepatitis from liver cancer related to  
XX cirrhosis; (7) screening for an agent capable of modulating the onset or  
XX progression of liver cancer; (8) a composition comprising at least two  
XX oligonucleotides comprising a sequence that specifically hybridizes to  
XX any of the genes; (9) a solid support comprising the at least two  
XX oligonucleotides; (10) a computer system comprising a database containing  
XX information identifying the level in liver tissue of a set of genes; (11)  
XX a method for using the computer system to present information identifying  
XX the expression level in tissue or cell of any of the genes; and (12) a  
XX therapeutic agent for slowing or halting the progression of liver cancer.  
XX The methods are useful for treating liver cancer associated with chronic  
XX hepatitis or cirrhosis. The present sequence represents a specifically



```
QY 660 aGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLe 680
Db 1758 AGGAACCTGGTATTCGGTGTGTGTGAACACACAGGGTGGTCTCAGTGTATCTCTGGTGGGGCT 1817
QY 680 uAlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAs 700
Db 1818 GGCACTCATGACACAGAGAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGA 1877
QY 700 pHisAspArgCysAspGlnHisIsthAspCysTySerCysThrAlaAsnThrAsnAspCy 720
Db 1878 CCATGACAGATGTGACACGACACACAGATTGTATACAGCTGCACAGCAACCAATGACTG 1937
QY 720 sHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSe 740
Db 1938 CCACCTGGTGCATGACCATGCTGTTGTTCCCGAGAACCCACAGCTGCTCAGAGGCCAGATCTC 1997
QY 740 rIlePheArgTyTrpGluAsnCysProLysAspAsnProMetTyTrpCysAsnLysLysTh 760
Db 1998 CATTGTTAGGTATGAGATTGCCCAAGGATAACCCATATGTACTACTGTAAACCAAGAGAC 2057
QY 760 rSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCy 780
Db 2058 CAGCTGCGAGAGCTGTGCGCTGGACCAAGACTGCCAGTGGAGGCCCGGAATCAGGAGTG 2117
QY 780 sIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLe 800
Db 2118 CATTGCCCTGCCGGAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAACTCATGTTT 2177
QY 800 uLysIleThrThrAlaLysGluAsnTyAspAsnAlaLysLeuPheCysArgAsnHisAs 820
Db 2178 GAAATTAATCTACTGCGCAAGGAGATTATGACAACTGCTAAATGTTTCTGTAGGAACCCACAA 2237
QY 820 nAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuAr 840
Db 2238 TGCCCTTTTGGCTTCTCTTACAACCCAGAGAGTAGAATTTGTCTTAAAGCAGCTGCG 2297
QY 840 gIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuAr 860
Db 2298 AATAATGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTGGCCCTTCG 2357
QY 860 gLysIleAsnValSerTyTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLe 880
Db 2358 GAAGATCAATGTGCTCTACTGGTGGGAGAGATGTGCCCATTTTACAAATAGTTTACT 2417
QY 880 uGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSe 900
Db 2418 ACAGTGGATGCCGTCTGAGCCAGTGAAGCTGGATTCTGTGGAATTTATCAGAACCCAG 2477
QY 900 rThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluAr 920
Db 2478 TACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAG 2537
QY 920 gProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGl 940
Db 2538 GCCTGCAAAACNACAGTGTAGACAGTGGCGGACACCATGTCCCTTGAGACAGCATGTGG 2597
QY 940 yAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAs 960
Db 2598 AGATTGCAACCGAGCGGCGAGCTCTGAGTCATGTGGTGACAGCAACATGAAGCAGTGTGGGA 2657
QY 960 pSerAsnAlaTyValAlaSerPheProPheGlyClnCysMetGlnTrpTyThrMetSe 980
Db 2658 CTCCTAATGCCATGTGGGCTCCTCCCTTTTGGCCAGGTATGAATGGTATACGATGAG 2717
QY 980 rThrCysProProGluAsnCysSerGlyTyCysThrCysSerHisCysLeuGluGlnPr 1000
Db 2718 CACCTGCCCTCCCTGAAATGTTGAGGCTACTGTACTGTACTGTAGTCATTGCTGGAGCAACC 2777
QY 1000 oGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySe 1020
Db 2778 AGGCTGTGGCTGGTACTGATCCCGCAATACTGCGCAAGGAAATGCATAGAGGGGTC 2837
```

```
QY 1020 rTyLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyTrpGlnPr 1040
Db 2838 CTATAAGGACCACTAGTGAAGATGCCCTTCGAAGCCCTACAGGAAATTTCTATCCACAGCC 2897
QY 1040 oLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyTrpAsnTrpSerPheIleHisCy 1060
Db 2898 CCTGCTCAATTCAGCATGTGCTCAGAGCAGCAGCATACAACCTGCTCTTCACTCACTG 2957
QY 1060 sProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCy 1080
Db 2958 TCCAGCTTGCATATGCAAGCGGCACAGTAATGCAATCAATCAGAGCATCTGTGAGAGTG 3017
QY 1080 sGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyTrpGlyAspPr 1100
Db 3018 TGAGAACCTTGACACAGGCAAGCACTGCGAGCTGCATATCTGTGGTTTCACTGGTGATCC 3077
QY 1100 oThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnTh 1120
Db 3078 CACCAATGAGGGAATGTGAGCATGCAAGTGCATGCGCAGCGCTCTCTGTGCAACAC 3137
QY 1120 rAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCy 1140
Db 3138 CAACACGGGCAAGTGTCTTGACACCAAGGCGGTCAAGGGGCGAGGTGCCAGCTATG 3197
QY 1140 sGluValGluAsnArgTyGlnGlyAsnProLeuArgGlyThrCysTyTrpThrLeuLe 1160
Db 3198 TGAGGTAGAAATCGATACCAAGAAACCTCTCAGAGGAACATGTTATTACTCTCT 3257
QY 1160 uIleAspTyTrpGlnPheThrPheSerLeuSerGlnGluAspAspArgTyTrpThrAlaI 1180
Db 3258 TATGTACTATCAGTTACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTAT 3317
QY 1180 eAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSe 1200
Db 3318 CAATTTTGGCTACTCTCTGACGAACAAACAGGATTTGGACATGTTTCAATCAATGCCCTC 3377
QY 1200 rLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGl 1220
Db 3378 CAAGAATTTCAACCTCAACATCACCTGGGCTGCGAGTTTCTCAGCTGGAAACCCAGGCTGG 3437
QY 1220 yGluGluMetProValValSerLysThrAsnIleLysGluTyLysAspSerPheSerAs 1240
Db 3438 AGAAGAGATGCCCTGTGTTTCAAAACCAACCATTAAGGAGTACAAGATAGTTTCTCTAA 3497
QY 1240 nGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyValSerAsnPh 1260
Db 3498 TGAGAAGTTTGAATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTGAGTAATTT 3557
QY 1260 eThrTrpProIleLysIleGlnIleAlaPheSerGln 1272
Db 3558 CACCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3594
RESULT 12
AAS72660
ID AAS72660 standard; cDNA; 3490 BP.
XX AAS72660;
DT 13-FEB-2002 (first entry)
XX
XX
DE DNA encoding novel human diagnostic protein #8464.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX OS
XX WO200175067-A2.
XX
XX 11-OCT-2001.
XX
XX 30-MAR-2001; 2001WO-US008631.
XX
XX
```

PR 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 PA (HYSE-) HYSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR P-PSDB; ABG08473.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 8464; 103pp; English.  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activities. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	0	Length:	3490
Score:	5622.00	Matches:	1070
Percent Similarity:	92.77%	Conservative:	8
Best Local Similarity:	92.08%	Mismatches:	18
Query Match:	71.03%	Indels:	67
DB:	5	Gaps:	6

US-09-787-097-12 (1-1429) x AAS72660 (1-3490)

QY	137	ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr	156
DB	22	AGACTAACTGGATCTCTGGTGTGACAGATGACCTGGAAATATAAATACAAAACG	81
QY	157	LysCysThrTrpLeuIleGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis	176
DB	82	AAGTGCACGTGGCTCATTTGAAGCAGCAGCAATAGAAATAGAGACTTCGTTTCATCAT	141
QY	177	PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla	196
DB	142	TTTGCTACAGATGTAGTTGGACCATTTATATGTTTATGATGGGACCTCAATTTATGCA	201
QY	197	ProLeuValAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal	216
DB	202	CCGCTAGTTGCTGCATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTC	261
QY	217	ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr	236
DB	262	CCTGAGTTGTTGCCACATCAGTTATGCTTGTGCTGCTATTTTATGATGCTGCTTAT	321
QY	237	AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly	256

DB	322	AAATTGACTGGATTTAATATATTACTTACAGTGTGTGATATGTGTCCAAATAAATGCTCAGGC	381
QY	257	ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn	276
DB	382	CGAGGAGAGTGAAGATCAGTAATAGCAGCGATCTGTTGAATGGAATGTTCTGAAAC	441
QY	277	TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg	296
DB	442	TGGAAGAGTGAAGATGTCATCTCTCACTGACAGACAACTGGTTTCTCTCATCGA	501
QY	297	GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro	316
DB	502	GGCATCTGCAATTCAGTGTGACAGGATGCTCTCTCTCAGACTGGCAGGCTCT	561
QY	317	GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAs	336
DB	562	GGATGTTTCACTTCTGTACAGCTAACCACTGTCATTTGGGACTCGAGAGATATCTAA	621
QY	336	n-LeuLysLeu-ProArgAlaSerHisLysAlaValValAsnGlnAsnIleValTrpVal	355
DB	622	CTTTAAAGCTCCCCAGAGCATCTCATAAAGCTGTGTCATATGGAACATTTATGGT	681
QY	356	ValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAla	375
DB	682	GTTGGAGGATATATGTTCAACCACTCAGATTATACATGTTCTAGCGTATGACCTTGT	741
QY	376	SerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHis	395
DB	742	TCTAGGAGTGGCTTCCACTTAACCGTTCTGTCAACAATGTGGTTAGATATGTCAT	801
QY	396	SerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyLysIleAspProThrGly	415
DB	802	TCCTTGGCATTTACAAGGATATAATTTACATGTATGGAGGAAATTAATCAACTGGG	861
QY	416	AsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThr	435
DB	862	AATGTACCCANAGATTGAGATTTTTCACATTCATATAGTCATGCGGTGTTGTGACC	921
QY	436	ProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLys	455
DB	922	CCTAAGCCAAAGAGCAGTATGCACTGGTGGGCACTCTGCACACATTTTACACTGAAG	981
QY	456	AsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSer	475
DB	982	AATGGCCGAGTGTGTCATGTCGTCATCTTGGTCACTGCTCTCTATGATATATAAGC	1041
QY	476	AsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAla	495
DB	1042	AATGTGCAGGATATGATTTGGATAAGAACACATGAGTATATATACACACCCAGGCTGC	1101
QY	496	LeuValGlnGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyr	515
DB	1102	CTTGTGCAAGGGGTTCAGCCCATAGCAGTGTTCAGACCATAGGACAGGCGCTATAC	1161
QY	516	ValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspLeuTyr	535
DB	1162	GTTTCATGGTGTACAGGCTTTTCAGTGCCTAATAGTACCGGCTTCAGATCTCTAC	1221
QY	536	ArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPheArgTyr	555
DB	1222	CGATATGATGTGATACCCAGATGGACCATTTCTTAAAGCAGACGCGATTTTCCGTTAC	1281
QY	556	LeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsn	575
DB	1282	TTGCACACAGCTGTGATAGTGTGACCATGCTGTGTTGGAGGAAACACACAAAT	1341
QY	576	AspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIle	595
DB	1342	GACACATCTATGAGCCATGGCGCAATGCTTCTCTCAGATTTTCATGCGCTATGACAT	1401
QY	596	AlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPhe	615
DB	1402	GCCTGTGACCGCTGTGTCAGTCTCCAGACCTGATCTCCACCATGATGTCACAGATTT	1461



```
QY 616 GlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeu 635
Db 1462 GGCCATTGAGGAGTCTTACACACAGACCATGATGATGTCGGTGGTTCAATAGTCTC 1521
QY 636 LeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAla 655
Db 1522 CTCCTCAGCGACATCTCTGGTATTACCTCGGAAACAGTGTGATGCGCATCGGAGTGAAGCC 1581
QY 656 AlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCys 675
Db 1582 GCTTGTATTAGCAGCAGGACCTGGTATTCTGGTGTCT- 1616
QY 676 IleSerTrpAlaLeuAlaThrAspGluGlnGluGluLeuLysSerGluCysPheSer 695
Db 1616 ----- 1616
QY 696 LysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAla 715
Db 1617 -----CTTGACCATGACAGATGTGACACAGACACAGATTGTTACAGCTGCACAGCC 1667
QY 716 AsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSer 735
Db 1668 AACACCAATGACTGCTCACTGGTGTCAATGACCAATGTGTCTCCAGCAACCAACAGCTGCTCA 1727
QY 736 GluGlyGlnIleSerIlePheArg-TyrGluAsnCysProLysAspAsnProMetTyrTy 755
Db 1728 GAAGCC-----CAGTATGAGAAATGCTCCCAAGGATACCCCACTGACTA 1772
QY 755 rCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluPr 775
Db 1773 CTGTAAACAAGAAGACACAGCTGCAGAGCTGTGCTGACCCAGAACCTGCCAGTGGGAGCC 1832
QY 775 oArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuVa 795
Db 1833 CCGGAATCAGAGGTGCAATGCTCCCTGCCCGAAATATCTGTGGCAATGCTGGCAATTTGGT 1892
QY 795 lGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPh 815
Db 1893 TGGAAACTCACTGTTGAAATTTACTACTGCCAAGGAGAAATATGACAAATGCTAAATTTGT 1952
QY 815 eCysArgAsnHisAsnAlaLeuAlaSerLeuThrThrGlnLysLysVal-GluPheV 835
Db 1953 CTGTAGGAACACCAATGCCCTTTTGGCTTCTTACACCCAGCAAGAGGTAGGAATTTG 2012
QY 835 alLeuLys--GlnLeuArgIleMetGln--SerSerGlnSerMetSerLysLeuThrLe 853
Db 2013 TCCITTAAGCGAGGTGCGAATAATGCAAGTCACTCAGAGGCAATGCTCAAGGCTCACCTT 2072
QY 853 uThrProTrp---ValGlyLeuArgLysIleAsn-ValSerTyrTrpCysTrpGluAspM 872
Db 2073 AACCCCATGGGTGGGCTTGGGAGAGATCAATTTGTCTCTACTGTTGCTGGGAGATA 2132
QY 872 etSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyP 892
Db 2133 TGTCCCCAATTACAAATAGTTACTTACAGTGGATGCCGTCTGAGCCCCAGTATGCTGGAT 2192
QY 892 heCysGly---IleLeuSerGluProSer--ThrArgGlyLeuLysAlaAlaThrCysIl 910
Db 2193 TCTGTGGGAATTTATTTCAGGAACCCAGTTACTTGGGGACTGAAGGCTGCAACCTGCAT 2252
QY 910 e-AsnProLeu-AsnGlySerValCysGlu-ArgPro-AlaAsnHisSer-AlaLysGln 928
Db 2253 TCAACCCCATCTCAATGGTAGTGTCTGTGAAAGGCTGTGGCAACCAACAGTGGTTAAGCAG 2312
QY 929 CysArgThrPro---CysAlaLeuArg-ThrAlaCysGlyAspCysThrSerGlySerSe 947
Db 2313 TGCCCGACACCATGCTGTCCTTTGAGGACAGCATGTGAGATTGACACCGCGCAGCTC 2372
QY 947 tGluCys-MetTrpCysSerAsnMetLys-GlnCysValAspSerAsnAlaTyrValAla 966
Db 2373 TGAGTGGCATGTGGTGAGCAACAAGAAGCAGTGTGTGGACTGCTCAATGCTCAATGTGGCC 2432
```

```
QY 967 SerPheProPhe-GlyGlnCysMetGluTrpTyrThrMet-SerThrCysProProGluA 986
Db 2433 TCCTTCCCTTTGGGCGAGTGTATGGAATGGTATAGATGAGGCACCTCCCTCCCTGAAA 2492
QY 986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro-GlyCysGlyTrpCys 1005
Db 2493 ATTGTTACGCTACTACTCTAGTCTAGTCTGTTGGAGCAACAGGGCTGTGGCTGTGT 2552
QY 1006 ThrAspProSerAsnThrGlyLysGlyCysCysIleGluGlySerTyrLysGlyProVal 1025
Db 2553 ACTGATCCCAAGCAATACTGGCAAAAGGAAATGCATAGAGGGTTCCTATAAAGACCAAGT 2612
QY 1026 LysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSer 1045
Db 2613 AAGATCCCTTGGCAAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATCCAGC 2672
QY 1046 MetCysLeuGluAspSerArgTyrAsnTrp-SerPheIle-HisCysProAlaCysGlnC 1065
Db 2673 ATGTGCTAGAGACAGACAGATACAACTGGGTCTTTTCAATTCACCTGCTCCAGTTTGC 2732
QY 1065 ysAsn-GlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThr 1084
Db 2733 GGAACGGGGCCAGATGAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACC 2792
QY 1085 ThrGlyLysHisCysGluThrCysIleSerGlyPheTyr-GlyAspProThrAsn-GlyG 1104
Db 2793 ACAGGCAAGCACTGCGAGACCTGCATATCTGCTTCTACGGGTGATCCCAATTTGGAG 2852
QY 1104 lLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyL 1124
Db 2853 GGAATATGTGAGCCATCAAGTGGCAACGCGCTCTGTGCAACACCAACACCGGCA 2912
QY 1124 ysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluA 1144
Db 2913 AGTGCTTCTGCCACCACCAAGGGCGTCAAGGGGAGCGAGTGCACGCTATGTGAGGTAGA 2972
QY 1144 snArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrG 1164
Db 2973 ATCGATACCAAGGAAACCCCTCTCAGAGGAACATG-TATTATATCTCTCTTATTGACTATC 3032
QY 1164 lnPheThrPheSerLeuSerGlnGluAspAspArgTyrThrAlaIleAsnPheValA 1184
Db 3033 AGTTCACTTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGG 3092
QY 1184 laThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheA 1204
Db 3093 CTACTCTCGACGAACAAACACAGGATTTGACATGTTTCATCAATGCTCCAGAAATTTCA 3152
QY 1204 snLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluLysMetP 1224
Db 3153 ACCTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGC 3212
QY 1224 roValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheA 1244
Db 3213 CTGTTGTTCAAAACCAACATTAAGAGTACAAAGATAGTTTCTTANTAGAGAGTTTG 3272
QY 1244 spPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProI 1264
Db 3273 ATTTTGGCAACCAACCAATATACATTTCTTGTGTTTATGCTAGTATTTTCACTGCGCCA 3332
QY 1264 leLysIleGlnIle 1268
Db 3333 TCAAAATTCAGGTG 3346
RESULT 13
ABZ58709
ID ABZ58709 standard; cDNA; 4140 BP.
XX
AC ABZ58709;
XX
XX 14-APR-2003 (first entry)
XX
DE Human homologue of attractin/mahogany (HAM) polypeptide encoding cDNA.
```



XX HAM; homologue of attractin/mahogany; immunosuppressive; cytostatic;  
 KW antinflammatory; cardiant; osteopathic; gene therapy; human; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Location/Qualifiers  
 FH 1. 4140  
 FT /tag= a  
 FT /product= "HAM"  
 FT /note= "homologue of attractin/mahogany"  
 FT 1. 180  
 FT /tag= b  
 FT mat\_peptide 181. 4137  
 FT /tag= c  
 XX  
 XX WO200297120-A1.  
 XX  
 XX 05-DEC-2002.  
 XX  
 XX 23-MAY-2002; 2002WO-US016391.  
 XX  
 XX 25-MAY-2001; 2001US-0293608P.  
 XX 24-SEP-2001; 2001US-0324626P.  
 XX  
 XX (IMV ) IMMUNEX CORP.  
 XX  
 XX Anderson DN;  
 XX  
 XX WPI: 2003-140486/13.  
 XX P-PSDB; ABP71087.  
 XX  
 XX New Homologue of Attractin/Mahogany (HAM) polypeptide, useful for  
 PT treating HAM-associated disorder consisting of inflammatory, autoimmune,  
 PT cell proliferative or cardiovascular disorders.  
 XX  
 XX Claim 14; Fig 1A-D; 89pp; English.  
 XX  
 XX The invention relates to Homologue of Attractin/Mahogany (HAM)  
 CC polypeptides and encoding polynucleotides. The HAM polypeptides can be  
 CC expressed by standard recombinant methodology. The HAM polypeptides are  
 CC useful for treating HAM-associated disorder consisting of inflammatory,  
 CC autoimmune, graft-versus-host, neurological, myelination, cell  
 CC proliferative, cardiovascular, haematologic, liver, metabolic, weight or  
 CC bone disorder. The present sequence represents a cDNA encoding a human  
 CC HAM polypeptide  
 XX  
 XX Sequence 4140 BP; 1209 A; 816 C; 941 G; 1174 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 1 99e-263 Length: 4140  
 Score: 4637.50 Matches: 843  
 Percent Similarity: 73.88% Conservative: 209  
 Best Local Similarity: 59.20% Mismatches: 300  
 Query Match: 58.59% Indels: 72  
 DB: 7 Gaps: 23  
 US-09-787-097-12 (1-1429) x ABZ58709 (1-4140)  
 QY 28 GlyProHisTrpAspValThrArgAlaGlyArgProGly----- 42  
 DB 18 GGCCCGCACTG-----TACCCCGCAGCCAGCGCCCGGGGTGTGGAGGGCTCG 68  
 QY 43 LeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProLeuArgProArgLeuLeu 62  
 DB 69 GCCGCGGGCGCGCGGGGGGCGCTCTCTCGGCTGCTGCGGACACCTGGCT 128  
 QY 63 LeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuProCysGluAlaG1 82  
 DB 129 GCTGTGCTATGGCTT-----CCTCTACCTGGCGCTCTAC----- 162  
 QY 82 uAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAs 102  
 DB 1191 AAAAATCTCTCTGTTCTTGGACATGGTTCAGCAGTATGCTGTGGAGGACATTCAGCACA 1250

163 -----GGCAGGTGTCCAGTCCAGCCGTGCGA 191  
 QY 102 pArg-----ProCysValAsnGlyClyValArgCysAsnProGlyThrGlyGlnCysValCy 120  
 DB 192 GAGGACCGGCTCTGCTCTCG---GGCGCTGTGTCAACTCCACC-----TGCTCTG 242  
 QY 120 sProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyA-gPheArgLeuThrG1 140  
 DB 243 CGACCCGGGCTGGTGGGGACCCAGTGCAGCACTGCCAGGCGAGCTTCAAGTTAACAGA 302  
 QY 140 ySerSerGlyPheValThrAspGlyProGlyAsnTrpLysTrpLysCysThrTr 160  
 DB 303 ACCTTCTGGATATTAAACAGATGGCCCAATTAACTATAAATAAACTAAATGACTTG 362  
 QY 160 pLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrG1 180  
 DB 363 GCTCATTTGAAGCTATCCAAATGCAGTGTAAAGATTAAAGATTCAATCAATTTTGTACAGA 422  
 QY 180 uCysSerTrpAspHisLeuTrpValTrpAspGlyAspSerIleTrpAlaProLeuValAl 200  
 DB 423 ATGTAGCTGGATCATATGATGTTTATGATGAGATTCAATATATATGCACTTAAATAGC 482  
 QY 200 aAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVa 220  
 DB 483 TGTACTTAGTGGTTTGATAGTCCCTGAATAAGGGGCATGAACCTGTGCTGCTTAAATCTA 542  
 QY 220 lAlaThrSerGlyTrpAlaLeuLeuHisPhePheSerAspAlaAlaTrpAsnLeuThrG1 240  
 DB 543 TACTACATCTGGCTATGCACTGTATCATTTTATGATGCTGCTGCTTAAATCTA 602  
 QY 240 yPheAsnIleThrTrpSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGlyCy 260  
 DB 603 TTTCAACATTTTCTATTCAATCAATCTTGTCTTAACTTCTCTGTCATGGAAGTG 662  
 QY 260 sLysIleSerAsnSer---SerGluThrValGluCysGluCysSerGluAsnTrpLysG1 279  
 DB 663 TACAACCTAGTGTCTGTCTCCAAAGTCAAGTATATTGTGAATGTGATAAATACTGGAAGG 722  
 QY 279 yGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyLeCy 299  
 DB 723 TGAAGCTTGTGATATTCTTACTGTAAAGCCAAATTCGGCAGTCCAGATCAGCGTTACTG 782  
 QY 299 sAsnSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSe 319  
 DB 783 TGACCTGACTGGAGAAAATATTGCTCTGCAATGATAGTTCGCAAGGCTCTGATTGTT 842  
 QY 319 rValProValProAlaAsnGlnSerPheTrpThrArgGluGluTrpSerAsnLeuLys-- 338  
 DB 843 TTTGAATGTTCCCTCTACTGAGTCTTACTGGATT-----CTGCCAAACGTTTAAACC 893  
 QY 339 -----LeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTr 354  
 DB 894 CTTTCAGTCTTCTGTAGTTCGGGCTTCATATAAGCAGTTTACAGGGAAATTTATGTG 953  
 QY 354 pValValGlyGlyTrpMetPheAsnHisSerAspTrpAsnMetValLeuAlaTyAspLe 374  
 DB 954 GGTGATTGGTGATATACCTTTAACTACAGTCTCTTTCAATATGGTCTCTAAATTTACA 1013  
 QY 374 uAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTrG1 394  
 DB 1014 AGAAGCAGTATATG---AATGTAGGAACCTCCATCAAGGGACCTCTCCAGAGATATGG 1070  
 QY 394 yHisSerLeuAlaLeuTyLysAspLysIleTrpMetTrpGlyGlyLysIleAspProTh 414  
 DB 1071 ACACCTCTCTGTTTATATACAGAAAACATCTTTATGTATGGAGGACAGATTGAACAAA 1130  
 QY 414 r---GlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpVal 433  
 DB 1131 TGATGCAATATGCACAGATGAATTTATGGTGTTTTAAACATACATAGTCACTGATGAGAT 1190  
 QY 433 uLeuThrProLysAla-----LysGluGlnTrpAlaValValGlyHisSerAlaHi 450  
 DB 1191 AAAAATCTCTCTGTTCTTGGACATGGTTCAGCAGTATGCTGTGGAGGACATTCAGCACA 1250



[illegible]

## RESULT 14

RESOLUT 13  
ABZ58710

ID ABZ58710 standard; cDNA; 4134 BP.

XXXXXX

AC ABZ58710;

100

DT 14-APR-2003 (first entry)

1000

DE Murine homologue of attractin/mah

XX  
XX  
XX  
homologous  
of  
the  
variety/morph

KW HAM; homologue of attractin/mahog

KW antinflammatory; cardiant; osteo  
XXXX  
OS  
Miss SP.

```
FT      /product= "HAM"  
FT      /note= "homologue of attractin/mahogany"  
  
XX  
XX  
PN      WO200297120-A1.  
XX  
XX      05-DEC-2002.  
XX  
XX      23-MAY-2002; 2002WO-US016391.  
PF      XX  
XX      28-MAY-2001; 2001US-0293608P.  
PR      XX  
XX      24-SEP-2001; 2001US-0324628P.  
XX  
XX      (IMMV ) IMMUNEX CORP.  
PA      XX  
XX      Anderson DN;  
XX      PI  
XX      WPI; 2003-140486/13.  
DR      P-PSDB; ABF71102.
```

XX New Homologue of Attractin/Mahogany (HAM) polypeptide, useful for  
PT treating HAM-associated disorder consisting of inflammatory, autoimmune,  
PT cell proliferative or cardiovascular disorders.  
XX  
PS Claim 14: Fig 5A-D: 89pp: English.

Effect of interaction related to Homologous/Mahogany (HAM)

CC The invention relates to Homologue of Attractin/Mangony (HAM.)  
CC polypeptides and encoding polynucleotides. The HAM polypeptides can be

polypeptides and encoding polynucleotides. The HAM polypeptides can be expressed by standard recombinant methodology. The HAM polypeptides are

expressed by standard recombinant methodology. The HAM polypeptides are useful for treating HAM-associated disorder consisting of inflammatory,

autoimmune, graft-versus-host, neurological, myelination, cell

CC proliferative, cardiovascular, haematologic, liver, metabolic, weight or

CC bone disorder. The present sequence represents a cDNA encoding a mouse.

CC HAM polypeptide

100

SQ Sequence 4134 BP; 1085 A; 981 C; 1036 G; 1032 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.42e-262	Length:	4134
------------	-----------	---------	------

Score:	4623.00	Matches:	848
Concordance:	73.74%	Concordance:	70%

Percent Similarity:	73.74%	Conservative:	205
Post Local Similarity:	50.20%	Mismatches:	300

160 pLeulleGluGlyClnProAsnArgIleMetArgLeuArgPheAsnHisPhealaThrGI 180  
 360 GCTAATTGAAGGGTATCCAAATGCAGTGTAAAGGTTAAAGATTCAATCATTTTGCTACAGA 419  
 180 uCySerTrpAspHisLeuThrValTyrAspGlyAspSerIleTyrAlaProLeuValAl 200  
 420 ATGCAGCTGGGATCATATGATGTATGATGATGAGATTCTATATAGCACCCTTTAGTAGC 479  
 200 aAlaPheSerGlyLeulleValProGluuArgAspGlyAsnGluThrValProGluValVa 220  
 480 TGTACTTAGTGGCTTGATCGTTCCTGAAGTGAAGGGTAAACAGACGGTGCCTGAGGTGGT 539  
 220 lAlaThrSerGlyTyrAlaLeuLeuHisPheSerAspAlaalaTyrAsnLeuThrGI 240  
 540 CACAGCGTCTGGCTACCGCTGCTCCACTTTTTCAGCGATGCTGCATATAAACCCTAAGTAC 599  
 240 yPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCy 260  
 600 CTTCAACATTTTATTTCGATCAATTCTCTCCTAACCACTGCTCTGTCATCGAAGAGTG 659  
 260 sLysIleSerAsnSer---SerGluThrValGluCysGluCysSerGluAsnTrpLysGI 279  
 660 TACAACAGGTGCTCTGCTTCAGTCAAGTGTATTTGTGAATGGCACAATPACTGGAAAGG 719  
 279 YGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCy 299  
 720 GGAAGCATGTGACATCTCTTACTGTAAAGCCAAATTGTGGAGTCCAGATCATGGCTACTG 779  
 299 sAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSe 319  
 780 TGACCTAAACAGGAGAGAAACTCTGTGCTGCAACGATAGTTGGCAAGCCCGCAGATTGTTTC 839  
 319 rValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLys-- 338  
 840 TCTGAATGTCCTTCTACTGAGCTTACTGGATT-----TTGCCAAATGTTTAAACC 890  
 339 -----LeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTr 354  
 891 CTTCAGCCCTCCGTAGGTGGGCTTCACATAAGCAGTTTTCATCGGAAATTCATGTG 950  
 354 pValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValleualaTyrAspLe 374  
 951 GGTGATTGGAGGATATCTTTTAACTACAGTCTCTTTTCAAATGGTTCGAAATTACAATTT 1010  
 374 uAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnVal-----Va 390  
 1011 AGAAGCAGTATATGG-----ATGTAGGTGCTGTATCAAGGGCCCTCT 1055  
 390 lValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLy 410  
 1056 TCAGAGATACGGACATCTCTTGGCCCTCTATCAGAAACAACTCTTATGTATCGGAGGCAG 1115  
 410 sLLeAspProThr---GlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGI 429  
 1116 AATGGACACAGTATGATGCCACGTCACCGATGAATATGGGTGTTAAACGTACCGAGTCA 1175  
 429 uSerTrpValLeuLeuThrProLysAla-----LysGlnTyrAlaValValGI 446  
 1176 ATCATGGAGCAGCAAAACCCCACTGCTTCCTGGCCACAGTCAGCAGTACCGTGGAGGG 1235  
 446 yHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheGI 466  
 1236 ACACCTCGGCACACATCATGGAGCTGGACAGTAGACGTGGTTCATGATTCATATTATGG 1295  
 466 yHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnTh 486  
 1296 ATATTCTGCATATATATGGCTATACACAGCAGCATACAGAAATACATATCTCTCAACAC 1355  
 486 rTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSerSerVa 506  
 1356 TTGGCTAGTTCAGAAACGAAAGAGGCCATTGTGCAAGTGGATATGGCAGCACAAGTGT 1415

QY	506	lTyRAspHiSaRgThrArgAlaLeuTyrValHisGlyGlyTyrIlySaIaPheSerAlaAs	526
DB	1416	GTATGACGAAGTGAACCAAGTCCATCTATGTTTCACGGAGGCTACAAAGCATGCTCGCA	1475
QY	526	nLYSTyArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTTPThrIl	546
DB	1476	TAAGTACGGCTGTGTGGACGACCTCTATAGTACGAAGTCAACACCAAGGACTTGGACTAT	1535
QY	546	eLeuLYAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMe	566
DB	1536	TTTCAAGAAGTGGGTTCGCATACCTTCATCTAGCTGTTCATTATCAATGGAGCTAT	1595
QY	566	tLeuValPheGlyGLVAsnThrHisAsnAspThrSerMetSerHisGlyAlaIlyCysPh	586
DB	1596	GCTTATTTTGGAGAAATACCCATAATGATACTTCCCTGAGTAACGGTGCNAATGTTT	1655
QY	586	eSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTyrSerValLeuProArgPr	606
DB	1656	TTCTGCCGANTTCTCGCATGATGACATAGCTTGTGACGAATGGAAGACATTCCTAAACC	1715
QY	606	aAspLeuHiHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMe	626
DB	1716	AAATCTCCATAGACAGCTCAACCATTTGGGCATTTCTGCAGTCGTCAATCAATGGGTCAAT	1775
QY	626	tTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGl	646
DB	1776	GTATATATTGGTGGATTTTCTAGTGACTCCITTAATGATATCTTGTGTATAAACCCCC	1835
QY	646	uGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCys	666
DB	1836	AAATTGCAAGCTTTCCGAGATGAAGAACTGTGCAGAAACGCTGGTCCAGGGATTAATATG	1895
QY	666	sValTIPAsnThrGlySerGlnCysIleSerTIPAlaLeuAlaThrAspGluGlnGl	686
DB	1896	TGTTTGGAAAT-----AGAAATCACTGTGAATCTTTGG-----GAGTCTGGGAATACAA	1943
QY	686	uGluLYsLeuLYsSerGluCysPheSerLYsArgThrLeuAspHisAspArgCysAspGl	706
DB	1944	TAATATTCTCAGAGCCAAAGTCCCTCCACAGACAGCTGCTACCCGATGACAGATTTACAG	2003
QY	706	nHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisnTPCysAsnAsp--	725
DB	2004	ATATGTGACTGTGCCAGCTGCACAGCCACACAGCGGTGCCAGTGGTGTGACAGAA	2063
QY	726	-HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGl	745
DB	2064	GAAATGCATCTCAGCCAGCAGCAACTGCAGC-----ACGTCTGTTCAGAAACTACAC	2114
QY	745	uAsnCysProLYsAspAsnProMetTyrTyrCysAsnLYsLYsThrSerCysArgSerCys	765
DB	2115	TAAATGTCATATAAGAAATGAGCAG---ATTGTGTAAACAACTTACAGAGCTGTAAAGCTG	2171
QY	765	sAlaLeuAspGlnAsnCysGlnTTPGluProArgAsnGlnGluCysIleAlaLeuProGl	785
DB	2172	TTCACTCAACTTGAATTGCCAGTGGATCAGCGCAGCAGGAATGTCAGGCTTTCAGCAGC	2231
QY	785	uAsnIleCysGlyIleGlyTIPHisLeuValGlyAsnSerCysLeuLYsIleThrThrAl	805
DB	2232	TCACCTTGTGGAAGCGCTGGAATCATGTGTGGGAGCGCTGTCTTCGAATCAATTCACG	2291
QY	805	LYsGluAsnTyrAspAsnAlaLYsLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSe	825
DB	2292	TCGAGAAGCTATGATTAATGCCAAACTTTATTTGCTATAACTCAGTGGAAATCTCGCTC	2351
QY	825	rLeuThrThrGlnLYsLYsValGluPheValLeuLYsValLeuArgIleMetGlnSerSe	845
DB	2352	CCTGCACACGTCACAGGAGGTGGATTTGTGTGATGAATA-----	2394
QY	845	rGlnSerMetSerLYsLeuThrLeuThrProTTPValGlyLeuArgLYsIleAsnValSe	865
DB	2395	-CAGAAGTTTCACAGCAGAAAGTGTCCCGTGGGTAGGCGCTACCGAAGATCAACATATC	2453
QY	865	rTyrTIPCysTProGluAspMetSerProPheThrAsnSerLeuLeuGlnTIProMetProSe	885

Db 2454 CTACTGGGATGGAGGAGCATGCTCTCTTCAATAACAGCCCTGAGTGGCTTCCTGG 2513  
 QY 885 rGluProSerAspAlaGlyPheCysGlyLeuSerGluProSerThrArgGlyLeuLy 905  
 Db 2514 TGAGCCAAATGACTCTGGATTCTGTCCTACTTAGAAGGGCTGCAGTGGCAGGGTTAAA 2573  
 QY 905 sAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAla----- 922  
 Db 2574 AGCAAAACCTTGCAATCCATGCCAGATGGACTCGCTTTGTGAAAGCCTGTAGTAAGCCC 2633  
 QY 923 -AsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCy 942  
 Db 2634 AAATCAGAACGGAGCGCTGCAAGAGCGTCTCCCTGAGGACCTCTCGGCCAACTG 2693  
 QY 942 sThrSerGlySerSerGluCysMetTrpCysSerAsnMetIysGlnCysValAspSerAs 962  
 Db 2694 CACGAGCAGCGCATGCGAGTGCATGTGGTGACAGCAGCAGAGCGCTGTGGACTCAA 2753  
 QY 962 nAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrThrMetSerThrCy 982  
 Db 2754 CGCTTACATCATCTCTCTTCCCTACGACAGTGCCTGGAGTGGCAGACT---GCCACCTG 2810  
 QY 982 sPropoGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCy 1002  
 Db 2811 CTCGCTCAAAATTTCTGGGTAAAGAACCTGTGCAGACAGTGTGGAGCAGCCAGGGTG 2870  
 QY 1002 sGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLy 1022  
 Db 2871 TGGCTGGTCAACGATCTCTAGTAAACACAGGAAGGCTATTGTCATCGAAGGGTCTTCCG 2930  
 QY 1022 sGlyProValLysMetProSerGluAlaProThrGlyAsnPheTyrProGlnProLeuLe 1042  
 Db 2931 GGGCCCAATGAACATC-----GTGGGGTCCCAACACAGTGCAGTGGTTC 2975  
 QY 1042 uAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAl 1062  
 Db 2976 TGACACCAACCTCTGCCCCAGGAGAGAACTACGAGTGGTCTTTTATCCAGTGTCCAGC 3035  
 QY 1062 aCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAs 1082  
 Db 3036 TTGCCAGTGTATGACACACAGCAGTGCATCAACAAACACACGCTGTGTGAGCAGTGAAGA 3095  
 QY 1082 nLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAs 1102  
 Db 3096 TCTCACCCTGGCGACAGTGTGAGGAATGTCATGCCAGGTACTATGAGAGCCCAACCAA 3155  
 QY 1102 nGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnTh 1122  
 Db 3156 CGGTGGCGAGTGCACAGCTTGCAGCTGGCGGCCCATCGAACGCTGTCTCACCCTGCACAC 3215  
 QY 1122 rGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluVa 1142  
 Db 3216 GGGAAAGTCTTTTGCACACACAGGGGATCAAGGGTGACAGTGCAGCTATGTGACTC 3275  
 QY 1142 lGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAs 1162  
 Db 3276 TGAATAATCGCTATGTTGGTAAATCCATTAGGGGGACATGCTATTACAGTCTCTGATTGA 3335  
 QY 1162 pTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPh 1182  
 Db 3336 CTACCAAGTTTACTTTCAGCTTCTGAGGAAGATGACCGGACACACACTGCGCATCACTT 3395  
 QY 1182 eValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAs 1202  
 Db 3396 CATCGCCAAACCCAGAGCAGTCAACAAACCTGAGACATTCGATTAAATGCTTCCACAA 3455  
 QY 1202 nPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGl 1222  
 Db 3456 CTTTAAATCTCAACATTACGTGTCTAGTTGGCTCAACAGCTGAACCATATCTGGGAAGA 3515  
 QY 1222 uMetProValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLy 1242

Db 3516 GACTCTATAGTTTCTTAAGACAATAATAAGGAATACAGACAGCAGCTTTTCTTATCAAAA 3575  
 QY 1242 sPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTr 1262  
 Db 3576 ATTTAACTTCAGAAAGCAATCTTAATATCATCAATTTATGTGTATGTGAGCAACTTCTCTG 3635  
 QY 1262 pProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPh 1282  
 Db 3636 GCCTATTAAATAACATTCGCTTTTCAACAACAACACGATCATGATCTCTGTCAGATT 3695  
 QY 1282 ePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTyrIly 1302  
 Db 3696 CTTTGTCACTTCTCTCAGTTGTTTTTACTTACTTGTGGTGGCTGCTGGTCTGGAA 3755  
 QY 1302 sIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGl 1322  
 Db 3756 GATCAAAACAACCTTGTGGCTTCTCGTGGAGGGAGCAACTGCTTCGAGAAACGACAGCA 3815  
 QY 1322 nMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProPr 1342  
 Db 3816 GATGGCCAGCCCTCCCTTGTCTTCTGATGTAGCCCTGGAGTAGGAGCTGAACAGAC 3875  
 QY 1342 oAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPh 1362  
 Db 3876 AGACTTTCTCGAGGGCCCATTAGAGGGTCCCTTAAGCCATAGCCATCGAACCCCTGCC 3935  
 QY 1362 eGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyIy 1382  
 Db 3936 TGGGAACAGAGCTGCTGCTGACTGTGTTCTCTCTACCGAGAGGATCTTCAGGCC 3995  
 QY 1382 eProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGl 1402  
 Db 3996 CCCACCCCTGGGCGAGTCCAGCTTCTATCGCCAGTGCCTGATAGACATCTCACAGCA 4055  
 QY 1402 nMetProIleValTyrLysGlySerGlyAlaValArgAsnArgLysGlnGlnProPr 1422  
 Db 4056 GAAGCCCTCTGATTAATAAGACAAGACTTCTGGAGTCCGCAATCGGAAGCAGCTCTCCAC 4115  
 QY 1422 oAlaGlnProGlyThrCysIle 1429  
 Db 4116 ACGTCAA---GGAACTTGTGTC 4134  
 RESULT 15  
 AAZ91922  
 ID AAZ91922 standard; cDNA; 2625 BP.  
 XX  
 AC AAZ91922;  
 XX 08-JUN-2000 (first entry)  
 XX Human mahogany protein coding sequence #4.  
 XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
 XX weight regulation; cell therapy; body weight disorder; cachexia;  
 XX anorexia; hyperpigmentation; increased metabolic rate disorder;  
 XX hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
 XX Homo sapiens.  
 XX WO200005373-A2.  
 XX 03-FEB-2000.  
 XX 21-JUL-1999; 99WO-US016484.  
 XX 21-JUL-1998; 98US-0093630P.  
 XX 20-OCT-1998; 98US-0104978P.  
 XX 05-FEB-1999; 99US-00245041.  
 XX (MILL-) MILLENIUM PHARM INC.  
 XX Moore K, Nagle DL;  
 XX



```
QY 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrIleu 556
Db 1606 TATGATGTGGATACAGATGTGGACCATCTTCTTAAGGACAGCCGATTTTCCGTACTTG 1665
QY 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576
Db 1666 CACACAGCTGTGATAGTGGTGGACCATGCTGGTGTGGGGGAAACACACACATGAC 1725
QY 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596
Db 1726 ACATCTATCAGCCATGGCGCCAAATGCTTCTCTTCAGATTTTCATGGCCTATGACATTGCC 1785
QY 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616
Db 1786 TGTGACCGCTGTGTAGTGTCTCCAGACTGATCTCCACCATGATCTCCACAGATTGTCACAGATTGGC 1845
QY 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636
Db 1846 CATTGACGAGTCTTACACACAGCACCATGTATGTGTTCCGTGGTTTCAATAGTCTCCTC 1905
QY 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656
Db 1906 CTCAGCGCATCTGCTGATCTACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965
QY 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676
Db 1966 TGTGTAGCAGCAGGACCTGGTATTTCGGTGTGTGTGGAACACAGGGTGTCTCAGTGTATC 2025
QY 677 SerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696
Db 2026 TCGTGGCGCTGGCAACTGATGACACAGAGAGAAAAGTTAAAAATCAGAAATGTTTTCCAAA 2085
QY 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716
Db 2086 AGAATCTTTGACCATGACAGATGTGACCCAGCACACAGATTGTTACAGCTGCACAGCCAC 2145
QY 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736
Db 2146 ACCAATGACTGCCACTGTGTGCAATGACCATTTGTCTCCAGGAAACACACAGCTGTCTCAGAA 2205
QY 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756
Db 2206 GGCCAGATCTCCATTTTATAGGTATGAGAAATGCCCAAGGATAACCCCTATGTACTACTGT 2265
QY 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776
Db 2266 AACAGAGACACAGCTGCAGGAGCTGTGCCCTGGACACAGAACTGCCAGTGGGAGCCCGG 2325
QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCys-GlyIleGlyTrp 792
Db 2326 AATCAGGAGTGCATTGCCCTGCCCGGTAGGCGCTTGACGGGTCACTCTTGG 2374
```

Search completed: March 1, 2004, 02:00:36

Job time : 1389 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 1, 2004, 01:57:50 ; Search time 1024 Seconds  
(without alignments)  
5034.248 Million cell updates/sec

Title: US-09-787-097-12

Perfect score: 7915

Sequence: 1 MVAARAAAEALRRRTAATA.....KSGAVRNKQPPAQPTCI 1429

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5

Ygapop 10.0, Ygapext 0.5

Fgapext 6.0, Fgapext 7.0

Delop 6.0, Delext 7.0

Searched: 2353733 seqs, 180373377 residues

Total number of hits satisfying chosen parameters: 4707466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-Q=/cgn2\_1/USPTO.spool.p/US09787097/runat\_26022004\_062937\_23538/app\_query.fasta\_1.1607  
-DB=Published Applications NA -QFMT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -END=1 -END=1 -MATRIX=blom62  
-TRANS=human40 cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09787097\_@CGN\_1\_136\_@runat\_26022004\_062937\_23538  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA.\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
18: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
-----					

1	7696.5	97.2	8589	9	US-09-893-238-14	Sequence 14, Appl
2	7303	92.3	8827	9	US-09-893-238-1	Sequence 1, Appli
3	6870.5	86.8	4072	9	US-09-893-238-16	Sequence 16, Appl
4	4637	58.6	6733	16	US-10-197-824-1	Sequence 1, Appli
5	4136.5	52.3	2625	9	US-09-893-238-18	Sequence 18, Appl
6	3491.5	44.1	6370	9	US-09-893-238-12	Sequence 12, Appl
7	3435	43.4	2419	9	US-09-893-238-8	Sequence 8, Appli
8	1226.5	15.5	1051	9	US-09-893-238-10	Sequence 10, Appl
9	960.5	12.1	8589	15	US-10-085-198-15	Sequence 15, Appl
10	923	11.7	7244	15	US-10-120-988-143	Sequence 143, App
11	840	10.6	207433	15	US-10-277-216-5	Sequence 5, Appli
12	840	10.6	207433	16	US-10-126-022-5	Sequence 5, Appli
13	766	9.7	90050	9	US-09-893-238-5	Sequence 5, Appli
14	581.5	7.3	415	9	US-09-833-381-1916	Sequence 1916, Ap
15	536	6.8	625	14	US-10-198-846-5869	Sequence 5869, Ap
16	454	5.7	17056	9	US-09-893-238-3	Sequence 3, Appli
17	446	5.6	246	9	US-09-736-457-966	Sequence 966, App
18	446	5.6	246	9	US-09-902-941-966	Sequence 966, App
19	446	5.6	246	9	US-09-849-626-966	Sequence 966, App
20	446	5.6	246	14	US-10-017-754-966	Sequence 966, App
21	446	5.6	246	14	US-10-113-872-966	Sequence 966, App
22	431	5.4	531	9	US-09-864-761-15900	Sequence 15900, A
23	393	5.0	212	9	US-09-796-692-7610	Sequence 7610, Ap
24	393	5.0	212	14	US-10-040-862-7610	Sequence 7610, Ap
25	393	5.0	212	15	US-10-057-475B-7610	Sequence 7610, Ap
26	393	5.0	212	15	US-10-154-843B-7610	Sequence 7610, Ap
27	369.5	4.7	8084	14	US-10-004-113-56	Sequence 56, Appl
28	365	4.6	7596	14	US-10-004-113-57	Sequence 57, Appl
29	359	4.5	7693	12	US-10-072-012-133	Sequence 133, App
30	357	4.5	7332	9	US-09-944-849-1	Sequence 1, Appli
31	357	4.5	7673	9	US-09-815-925-1	Sequence 1, Appli
32	357	4.5	7673	15	US-10-294-006-1	Sequence 3, Appli
33	356.5	4.5	6728	15	US-10-190-115-3	Sequence 3, Appli
34	356.5	4.5	6728	15	US-10-369-072-3	Sequence 32404, A
35	340	4.3	193	9	US-09-864-761-32404	Sequence 1, Appli
36	340	4.3	7410	15	US-10-190-115-1	Sequence 1, Appli
37	340	4.3	7410	15	US-10-369-072-1	Sequence 59, Appl
38	334.5	4.2	7615	14	US-10-004-113-59	Sequence 169, App
39	334	4.2	5312	14	US-10-311-455-169	Sequence 21460, A
40	332	4.2	182	9	US-09-864-761-21460	Sequence 86, Appl
41	326.5	4.1	8091	14	US-10-101-510-86	Sequence 1, Appli
42	326.5	4.1	8091	14	US-10-356-625-1	Sequence 322, App
43	326.5	4.1	8091	15	US-10-159-563-322	Sequence 4720, Ap
44	326	4.1	375	9	US-09-864-761-4720	Sequence 133, App
45	303	3.8	11558	10	US-09-949-029-133	

#### ALIGNMENTS

#### RESULT 1

US-09-893-238-14  
; Sequence 14, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 8589  
; TYPE: DNA  
; ORGANISM: Homo sapiens



Db 1966 TGTTTAGCAGCAGGACCTGGTATTTCGGTGTGTGTGGAAACACAGGGTCTGTCTCAGTGTATC 2025  
Qy 677 SerTrrAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696  
Db 2026 TCGTGGGCGCTGGCAACTGATGAACAAGAAAGATTAAATCAGAAATGTTTTCAAA 2085  
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
Db 2086 AGAACTCTTGACCATGACAGATGTGACGACACACAGATTGTTACAGCTGCAGCCCAAC 2145  
Qy 717 ThrAsnAspCysHisTrrPAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
Db 2146 ACCAATGACTGCCATGTGTGATGACCAATTTGTCTCCCAAGAACCAAGCTGTCTGAAA 2205  
Qy 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
Db 2206 GCCCAGATCTCCATTTTGTAGTATGAGAAATTTGCCCAAGGATAACCTTATGTACTATGT 2265  
Qy 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrrPProArg 776  
Db 2266 AACAAAGAACCCAGCTGCAGGAGCTGTGCCCTGGACACAGAACTGCCAGTGGGAGCCCGG 2325  
Qy 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrrPHisLeuValGly 796  
Db 2326 AATCAGGAGTGCATTGCCCTGCCGAAATATCTGTGGCATTTGGTGGGA 2385  
Qy 797 AsnSerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
Db 2386 AACTCATGTTTGAATTTACTTACTGCCAAGGAGAAATATGACATGCTAAATTTGTTGT 2445  
Qy 817 ArgAsnHisAsnAlaLeuAlaSerLeuThrGlnLysLysValGluPheValLeu 836  
Db 2446 AGGAACCAATGCTTGTGCTTCTTAAACCCCAAGAAAGGTAGAAATTTGTGCTT 2505  
Qy 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrrP 856  
Db 2506 AAGCAGCTGGCAATTAATGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAAACCCCATGG 2565  
Qy 857 ValGlyLeuArgLysIleAsnValSerTyrTrrPTrpCysTrrPTrpGluAspMetSerProPheThr 876  
Db 2566 GTGGCCCTTCGGAAGATCAATGTGTCTTACTTGGTGTGGGAAGATATGTCCTCCATTTACA 2625  
Qy 877 AsnSerLeuLeuGlnTrrPMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
Db 2626 AATAGTTTACTACAGTGGATGGCGTCTGAGCCCAAGTGAATCTGTGGAAATTTTA 2685  
Qy 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
Db 2686 TCAGAACCCAGTACTCGGGGACTGAGGCTGCAACCTGCATCAACCCCACTCAATGGTAGT 2745  
Qy 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
Db 2746 GTCTGTGAAGGCCCTGCAACCCACAGTGTCTAAGCAGTCCCGACACCATGTGCCTTGAGG 2805  
Qy 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrrPTrpCysSerAsnMetLys 956  
Db 2806 ACAGCATGTGGAGATTGCACAGCGGCAGCTCTGAGTGCATGTGGTGAGCAACATGAAG 2865  
Qy 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrrP 976  
Db 2866 CAGTGTGTGACTCCAAATGCCATGTGGCTCTTCCCTTTTGGCCAGTGTATGGAATGG 2925  
Qy 977 TyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
Db 2926 TATACGATGAGCACCTGCCCTCCCTGAAATTTGTTGAGGCTACTGTACCTGTAGTCAATGTC 2985  
Qy 997 LeuGluGlnProGlyCysGlyTrrPTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1016  
Db 2986 TTGGACCAACAGGCTGTGGCTGTGCTACTGATCCAGCAATACTGGCAAGGGAATGTC 3045  
Qy 1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036  
Db 3046 ATAGAGGGTTCCTATAAAGGACAGTGAAGATGCTTCCCTTCCAGAGCCCTTACAGGAAATTC 3105

Qy 1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrrPser 1056  
Db 3106 TATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACACTGGTCT 3165  
Qy 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
Db 3166 TTCAATCACTGTCCAGCTTGCATGCAACGCGCCACAGTAAATGTCATCAATCAGAGCATC 3225  
Qy 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096  
Db 3226 TGTGAGAGTGTGAGAACTTGACCAAGCAAGCACTGCGAGACCTGCAATATCTGGCTTC 3285  
Qy 1097 TyrGlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116  
Db 3286 TAGCGTGTATCCCAACCAATGAGGGAATGTCAGCCATGCAAGTGAATGGCACCGGTCT 3345  
Qy 1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136  
Db 3346 CTGTGAAACACCAACACGGGCAAGTCTTCTGCACCACCAAGGCGTCAAGGGGACGAG 3405  
Qy 1137 CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr 1156  
Db 3406 TGCAGCTATGTAGGTAGAAATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTAT 3465  
Qy 1157 TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr 1176  
Db 3466 TATACTCTTCTTATTGACTATCAGTTTCACTTTAGCTTATCCAGGAAGATGATCGCTAT 3525  
Qy 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196  
Db 3526 TACACAGCTATCAATTTTGTGGTACTCTCTGACGACAAACACAGGATTTGGACATGTC 3585  
Qy 1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrrPAlaAlaSerPheSerAlaGly 1216  
Db 3586 ATCAATGCTCTCCAAAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGA 3645  
Qy 1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTrrLysAsp 1236  
Db 3646 ACCAGCGCTGGAAGAGATGCTGTGTGTTTCAAAACCAACCAATTAAGAGTACAAAGAT 3705  
Qy 1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256  
Db 3706 AGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCAACCAATATCACTTTCTTGTATTAT 3765  
Qy 1257 ValSerAsnPheThrTrrPProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276  
Db 3766 GTCAGTAATTTCACTTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCACAGCAATTTT 3825  
Qy 1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuVal 1296  
Db 3826 ATGGACCTGTGTACAGTTCTCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3885  
Qy 1297 AlaAlaValValTrrPlyLysGlnSerCysTrrPAlaSerArgArgGluGlnLeu 1316  
Db 3886 GCTGCTGTGTTTGGGAAGATCAAAAGATTTGTTGGGCTTCCAGAGCTAGAGAGCAACTT 3945  
Qy 1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336  
Db 3946 CTTGAGAGATGCAACAGATGGCCGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 4005  
Qy 1337 ThrAspGluGluProProAspLeuIleGlyLysSerIleLysThrValProLysProIle 1356  
Db 4006 ACAGATGAGGAGCTCTCTGATCTTATTGGGGGAGTATAAAGACTGTTCACAAACCATTT 4065  
Qy 1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376  
Db 4066 GCATGTGAGCGGTGTTTGGCAACAAAGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCT 4125  
Qy 1377 ArgGlyLeuGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396  
Db 4126 CGAGCGCTGTGGTGGCATCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4185



Db 1559 GGTATTGTTAACTCCGAAGCTAAGGATCATGATGAGTGGTTGGACACTCAGCACACAT 1618  
Qy 451 eValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTy 471  
Db 1619 TGTACATCGCATCTGCGCGTGTGTCATGTTGGTCACTCTCGGTCACTGCCACACTA 1678  
Qy 471 rGlyTyrlleSerAsnValGlnGluTyAspLeuAspLysAsnThrTrpSerIleLeuHi 491  
Db 1679 TGGATATATAAGCGTTGTGAGGAATATGACTTGGAAAAGAACACATGGAGTATATTACA 1738  
Qy 491 sThrGlnGlyAlaLeuValGlnGlyGlyTyGlyHisSerSerValTyAspHisArgTh 511  
Db 1739 TACTCAGGGTCTCTTGTGCAAGGGGGTTATGGCCACAGTAGTGTATATGATGACAGGAC 1798  
Qy 511 rArgAlaLeuTyValHisGlyGlyTyLysAlaPheSerAlaAsnLysTyArgLeuAl 531  
Db 1799 CAAGGCTCTGACGTTTCATGTTGCTACAGGCTTTCAGCGCCAAACAAATACCGGCTTC 1858  
Qy 531 aAspAspLeuTyArgTyAspValAspThrGlnMetTrpThrIleLeuLysAspSerAr 551  
Db 1859 AGATGACCTTACAGATACGATGTGGATACATGATGTGGACCAATCTTAAAGGACAGCG 1918  
Qy 551 gPhePheArgTyLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyG 571  
Db 1919 ATTTTTCGTTACTTGCATACAGCTGTGATGTAGTGGTGAACCAATGCTGTGTGTGGAG 1978  
Qy 571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591  
Db 1979 GAACACACACAATGACACTTCATGAGCCACGGTGCCAAATGCTCTCTCTCGGACTTCAT 2038  
Qy 591 tAlaTyAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAs 611  
Db 2039 GGCATTATGACATGCTTGTGACCGATGTCTAGTGTCTCCAGAGCTGAGCTCCATCATGA 2098  
Qy 611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyValPheGlyG 631  
Db 2099 TGTCAACAGATTGGCCATTGAGCATCTGTACAAACAGCACCATGATGTGTTCGGCGG 2158  
Qy 631 yPheAsnSerLeuLeuSerAspIleLeuValPheThrSerGlnGlnCysAspAlaHi 651  
Db 2159 CTTCAACAGCTCTCTCTCAGTGAGCTGTGCTTACCTCGAGCAGCTGCGATGCACA 2218  
Qy 651 sArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGl 671  
Db 2219 CCGAGTGAAGCTGCTGTGTGGCAGCAGGACCTGGTATCCGGTGTCTGTGGGACACACA 2278  
Qy 671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGlnGlnGluLysLeuLysSe 691  
Db 2279 GTGGTCTCGATGCTCTCTCGGAGTTGGCAACTGAAGAACAAAGTAAAGTTAAATC 2338  
Qy 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711  
Db 2339 AGAGTGTCTTCTTAAAGAACCTTGAACCATGACATGACATGATGATCACTGTGCTGTGAA 2398  
Qy 711 rSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAs 731  
Db 2399 CAGCTGCACAGCCATACCAATGACTGCCACTGGTGCAATGATCACTGTGCTGTGAA 2458  
Qy 731 nHisSerCysSerGluGlyGlnIleSerIlePheArgTyGlnAsnCysProLysAspAs 751  
Db 2459 CCACAGCTGCACAGAGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCAAGGATAA 2518  
Qy 751 nProMetTyTyCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCy 771  
Db 2519 CCCCATGTACTCTCATATAGAAACCCAGCTGCAGGAGCTGTGCCCTAGACCCAGACTG 2578  
Qy 771 sGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGl 791  
Db 2579 CCAGTGGGAGCCCGGAATCAAGAGTGCATCGCCCTGCCGGAATAATATCTGTGCAATGG 2638  
Qy 791 yTrpHisLeuValGlyAsnSerCysLeuLysIleThrAlaLysGluAsnTyAspAs 811  
Db 2639 CTGGCATTTGGTTGGAACCTGTGTCTGAAATCACTACTGTAAGGAGAAATTATGACA 2698

Qy 811 nAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysTy 831  
Db 2699 TGTCTAAATTTGCTCTGTAGGAACCAACAATGCCCTTTTGGTTCCCTCACATCCAGAGAA 2758  
Qy 831 sValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLe 851  
Db 2759 GGTGAGTTTGTCTTAAAGCAGCTTCGATTAATGCAATCATCTCAAGTATGTCCAAGCT 2818  
Qy 851 uThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyTrpCysTrpGluAs 871  
Db 2819 CATCTGACTCCCAAGGTTGGTCTTCGGAAGATCAATGTCTTACTGTGTGGGAGGA 2878  
Qy 871 pMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGl 891  
Db 2879 TATGTCTCCATTACAAATAGTTTGTGAGTGGATGCCATCTGAGCCAGTGTCTGTG 2938  
Qy 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911  
Db 2939 CTTCTGTGGATCTTCTGAGCCTAGTACTCGGGATTTAAAGGCTGCAACCTGCAATCAA 2998  
Qy 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgTh 931  
Db 2999 CCCTCTCAATGGCAGCGCTCTGTGAAGGCTGCACACACAGTGCACAGCAGTCCGGAC 3058  
Qy 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951  
Db 3059 ACCATGTGCCCTCGGACAGCGTGTGCGAGTGCACACTAGCAGCAGCTCGAGTGCATGTG 3118  
Qy 951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyValAlaSerPheProPheGl 971  
Db 3119 GTGCAGTACATGACAGCAGTGTGTGACCTCCAAATGCTACGTAGTGGCTCTTCCCTTTGG 3178  
Qy 971 yGlnCysMetGluTrpTyThrMetSerThrCysProProGluAsnCysSerGlyTyCy 991  
Db 3179 CCAGTGTATGGAATGTTATACGATGAGCAGCTGCCACCTGAAATTTGCTCTGGTACTG 3238  
Qy 991 sThrCysSerHisCysLeuGluClnProGlyCysGlyTyProCysThrAspProSerAsnTh 1011  
Db 3239 TACTCGAGCCATTGTCTTGAGCAGCCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3298  
Qy 1011 rGlyLysGlyLysCysIleGluGlySerTyLysGlyProValLysMetProSerGlnAl 1031  
Db 3299 TGGGAAGGAAATGTTATTGAGGCGCAGCTATAAAGGACCTGTGAAGATGCCGTGCACAGC 3358  
Qy 1031 aProThrGlyAsnPheTyProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSe 1051  
Db 3359 CTCTGAGGAAATGTGTATCCAGCCCTTCTGAACTCCAGCATGTGTAGAGGACAG 3418  
Qy 1051 rArgTyAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCy 1071  
Db 3419 CAGATCAACTGTGCTTTTATTCACTGTCCAGTTCAGCTGCCAGTGCACCGACACAGCAATG 3478  
Qy 1071 sIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrGlyLysHisCysGluTh 1091  
Db 3479 CATCAACAGAGTATCTGTGAGAAGTGTGAGGACCTGACACCGGCGAGCAGTCCGAGAC 3538  
Qy 1091 rCysIleSerGlyPheTyArgTyProThrAsnGlyGlyLysCysGlnProCysLysCy 1111  
Db 3539 CTGCATATCTGGTCTTCTATGGTGACCCGACTAATGGAGGCAATGTCACCCATGCAAGTG 3598  
Qy 1111 sAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGl 1131  
Db 3599 CAATGGCGCAGCATCTGTGCAACACCAACACCGGCAAGTGTCTCTGTATCCACCAAGG 3658  
Qy 1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyThrGlnGlyAsnProLe 1151  
Db 3659 TGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAAAATCGATACCAAGAAACCTCT 3718  
Qy 1151 uArgGlyThrCysTyTrpThrLeuLeuIleAspTyThrGlnPheThrPheSerLeuSerGl 1171  
Db 3719 CAAAGGACATGCTACTATACCTTCTCATTTGACTATCAGTTTACCTTTAGCTGTCCCA 3778

1171	Qy	nGIuSPaSPArGTYrTYrThrAlaIleAsnPheValAlaThrProASPGLuGlnAsnAr	1191
3779	Db	GGAAAGCAGCGCTACTACACAGCATCAACTTTGTGGCTACTCTCATGAACAAAACAG	3838
1191	Qy	gASpLeuASpMetPheIleAsnLaserLysAsnPheAsnLeuAsnIleThrTPAlaAl	1211
3839	Db	GGATTGGACATGTTTCATCAATGCCTCCAAAACCTTCAACCTCAACATCACCTGGGCGAC	3898
1211	Qy	aSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIl	1231
3899	Db	CAGCTTCCCGCCGGAACCCAGACCTGGAGAAGAGGTGCTGTGTGTTTCAAAAACCAACAT	3958
1231	Qy	eLysGluTYrLYIAASPserPheSerAsnGluLYSPheASPheArgAsnHisProAsnIl	1251
3959	Db	CAAGGAATACAAAGATAGCTTCTAATGAGAAATTTGATTTTGGCAACCATCCAAACAT	4018
1251	Qy	eThrPhePheValTYrValSerAsnPheThrTPProIleLysIleGlnIleAlaPheSe	1271
4019	Db	CACCTTTCTTTGTTTATGTCTAGTATTTTCACTTGGCCCATCAAAATTCAGATTGCTTCTC	4078
1271	Qy	rGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLe	1291
4079	Db	CCAGCAGAGCAACTTCTATGGACCTGGTACAGTTCTTCGTGACTTCTTCAGTTGTTCCT	4138
1291	Qy	uSerLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerAr	1311
4139	Db	CTCGTCTCTTCTGGTGGCTGCAGTGGTCTCGAAGATCAAGCAGACGTGTTTGGGCATCCAG	4198
1311	Qy	gArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerVa	1331
4199	Db	GCGGAGAGAGCACTTCTTTCGGGAGATGCACAGATGGCCAGCCGCCCTTTCGTTCTGT	4258
1331	Qy	lAsnValAlaLeuGluThrAspGluGluProProASPLeuIleGlySerIleLysTh	1351
4259	Db	AAACGTTGCCTTGGAAACAGATCAGGAGCCCTCTCGATCTTATTGGGGGAGTATAAAGAC	4318
1351	Qy	rValProLYSProIleAlaLeuGluProCysPheGlyAsnLYsAlaAlaValLeuSerVa	1371
4319	Db	TGTTCCCAACCCCATTCATTCGGAGCCGTGTTTGGCAACAAGCCCGCTGTCTCTCTGT	4378
1371	Qy	lPheValArgLeuProArgGlyLeuGlyGlyIleProProProGlyGlnSerGlyLeuAl	1391
4379	Db	GTTTGTGAGGCTCCCTCAGGCCCTGGGTGGCATCCCTCTCTCTGGGCGAGTCAGGTCTTGC	4438
1391	Qy	aValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTYrLysGluLysSe	1411
4439	Db	TGTGGCCAGCGCCCTGTGTGACATTTCTCAGCAGATCCCATAGTGTACAGGAGAGATC	4498
1411	Qy	rGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle	1429
4499	Db	AGAGCGCTGAGAAACCGAAGCAGACGCCCTCTGCACGCTGGACCTGCATC	4553

### RESULT 3

```

US-09-893-238-16
; Sequence 16, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16

```



Db 826 TGGAAAGGTGAAGCATGTGACATTCCTCACTGTGTACAGACAACTGTGGTTTTCCCTCATCGA 885  
Qy 297 GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTyrGlnGlyPro 316  
Db 886 GGCATCTGCAATTCAGTGATGTACAGAGATGCTCCTCTTCTCAGACTGGCAGGGTCT 945  
Qy 317 GlyCysSerValProValProAlaAsnGlnSerPheTyrThrArgGluGluTyrSerAsn 336  
Db 946 GGATGTTTCTGATTCCTGTACCACTAACCAGATCATTTTGGACTCGAGAGGAATATCTTAAC 1005  
Qy 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTyrValVal 356  
Db 1006 TTAAGATCCCCAGAGCATCTCATAAAGCTGTGGTCAATGGAACATATATGGGTGTT 1065  
Qy 357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376  
Db 1066 GGAGGATATATGTTCAACCACTCAGATATATACATGTTCTAGCGTATGACCTTGCTTCT 1125  
Qy 377 ArgGluTyrLeuProLeuAsnArgSerValAsnValValAlaTyrGlyHisSer 396  
Db 1126 AGGGATGGCTTCCACTAAACCGTTCTGTGAACAATGGTGTGTAGATATGGTCAATCT 1185  
Qy 397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416  
Db 1186 TTGGCATTTATACAGGATAAAATTTACATGTATGGAGGAAAAATTTGATTCAACTGGGAAT 1245  
Qy 417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTyrValLeuLeuThrPro 436  
Db 1246 GTGACCAATGAGTTGAGAGTGTTCACATTCATATGATGATGGTGTGTTGTGACCCCT 1305  
Qy 437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456  
Db 1306 AAGCAAGGAGGACGATGATGAGTGTGGTGGTCTGACACATTTGTACACTGAAGAAT 1365  
Qy 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476  
Db 1366 GGCCAGTGGTCACTCGTCTATCTTGGTCACTGCCCCCTCTATGGATATATTAAGCAAT 1425  
Qy 477 ValGlnGluTyrAspLeuAspLysAsnThrTyrSerIleLeuHisThrGlnGlyAlaLeu 496  
Db 1426 GTGCAGGAAATGATTTGGATGAAGAACACATGGAGTATATTACACCCAGGGTGGCCCT 1485  
Qy 497 ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516  
Db 1486 GTCCAGGGGGTTAGGGCATAGCAGTGTATTACGACCATAGGACCCAGGGCCCTATACGTT 1545  
Qy 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArg 536  
Db 1546 CATGTGGCTACAGCTTTAGTCCCAATAGTACCGGCTTGACAGATGATCTCTACCGA 1605  
Qy 537 TyrAspValAspThrGlnMetTyrThrIleLeuLysAspSerArgPheArgTyrLeu 556  
Db 1606 TATGATGTGGATPACCCAGATGTGGACCATTTCTTAAGGACAGCCGATTTTCCGTTACTTG 1665  
Qy 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576  
Db 1666 CACACAGCTGTGATAGTAGTGAACCAATGCTGGTGTGGGGGAAACACACACAATGAC 1725  
Qy 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596  
Db 1726 ACATCTATGAGCCATGGCGCCAAATGCTTCTTCAGATTTTCATGGCCTATGACATTGCC 1785  
Qy 597 CysAspArgTyrSerValLeuProArgProAspLeuHisAspValAsnArgPheGly 616  
Db 1786 TGTGACCGTGTGTCAAGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTGGC 1845  
Qy 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
Db 1846 CATTCAGCAGTCTTACACACAGCACCATGATGTGTTCCGTGGTTCATATGATCTCTC 1905  
Qy 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
Db 1906 CTCAGCGACATCTCGTATTCACCTCGGAACAGTGTGATGCGCATCGGAGTGAAGCCGCT 1965

Qy 657 CysLeuAlaAlaGlyProGlyIleArgCysValTyrAsnThrGlySerSerGlnCysIle 676  
Db 1966 TGTTTAGCAGCAGGACCTGGTATTCGGTGTGTGTGGAACACAGGGTCTGCTCAGTGTATC 2025  
Qy 677 SerTyrAlaAlaLeuAlaThrAspGluGlnGluLeuLysLeuLysSerGluCysPheSerLys 696  
Db 2026 TCGTGGCGGTGGCAACTGATGAACAGAGAAAGTAAATTCAGATGTTTTTCCAAA 2085  
Qy 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
Db 2086 AGAATCTTGCACCATGACAGATGTGACACGACACACAGATTTGTACAGTGCACAGCCAAC 2145  
Qy 717 ThrAsnAspCysHisTyrCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
Db 2146 ACCAATGACTGGCACTGGTGCATGCAATGTTGTCCCGAGGAACACACAGCTGCTCAGAA 2205  
Qy 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
Db 2206 GGCCAGATCTCCATTTTAGGTATGAGATTTGCCCCAGGATTAACCTATGACTACTGT 2265  
Qy 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTyrGluProArg 776  
Db 2266 AACAAAGAACCACTGCAGGAGCTGTGCCCTGACACAGAACTGCCAGTGGAGCCCGG 2325  
Qy 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTyrHisLeuValGly 796  
Db 2326 ATTCAGAGTGCATTTGCCCTGCCGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGA 2385  
Qy 797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
Db 2386 AACTCATGTTTGAATAATTACTACTGCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGT 2445  
Qy 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836  
Db 2446 AGAACCACAAATCCCTTTTGGCTTCTTTACAAACCAGAAAGGTAGAAATTTGTCCTT 2505  
Qy 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTyr 856  
Db 2506 AAGCAGCTGGATATATGACGATCTCAGAGCATGTCCAGCTCACTTAACCCCATGG 2565  
Qy 857 ValGlyLeuArgLysIleAsnValSerTyrTyrCysTyrProLysPheMetSerProPheThr 876  
Db 2566 GTCGGCTTTCGGAAGATCAATGTGCTCTACTGTGTCTGGGAAGATATGTCCTCATTTTACA 2625  
Qy 877 AsnSerLeuLeuGlnTyrMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
Db 2626 AATAGTTTACTACAGTGGATGCCGCTGAGCCCGAGTGTGGATCTGTGGAAATTTTA 2685  
Qy 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
Db 2686 TCAGAACCCAGTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745  
Qy 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
Db 2746 GTCTGTGAAGGCTCCAAACACACAGTGTAAAGCAGTGCAGACCATGTGCTGCTTGAGG 2805  
Qy 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLys 956  
Db 2806 ACAGCATGTGGAGATTGCACCGCGAGCTCTGAGTGCATGTGTGTGCAGCAACATGAAG 2865  
Qy 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyr 976  
Db 2866 CAGTGTGTGACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAAATGG 2925  
Qy 977 TyrThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
Db 2926 TATACATGAGCAGCTGCCCTGAAAATTTGTTAGGGTACTGTACCTGTAGTCAATTCG 2985  
Qy 997 LeuGluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCys 1016  
Db 2986 TTGAGCAACAGCGCTGTGGTGTGTACTGTATCCAGCAATACTGGCAAGGAAATGC 3045



1017	QY	IleGluGlySerTyrIysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe	1036
3046	DB	ATAGAGGGTTCTTATAAAGACACAGTGAAGATGCCCTTCGCAAGCCCTACAGAAATTTTC	3105
1037	QY	TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer	1056
3106	DB	TATCCACAGCCCTGTCTCAATTCACGATGTGTCTAGAGGACAGCAGATACAACACTGGTCT	3165
1057	QY	PheIleHisCysProAlaCysGlnCysAsnGlyHisSerIysCysIleAsnGlnSerIle	1076
3166	DB	TTCAATTCACCTGCTCCAGCTTGCCCAATCAACGGGCCACAGTAAATTCGATCAATCAGAGCATC	3225
1077	QY	CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGlnThrCysIleSerGlyPhe	1096
3226	DB	TGTGAGAGTGTGAGAACCTGCACCAAGGCCAGCATGCGAGACCTGCATATCTGGCTTC	3285
1097	QY	TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysIysCysAsnGlyHisAlaSer	1116
3286	DB	TACGGTGTATCCCAACCAATGGAGGAAATGTGCAGCCATGCAAGTGCATGGGCACGGGTCT	3345
1117	QY	LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu	1136
3346	DB	CTGTGCAACCAACACGGGCAAGTGCTTCTGCACCAAGGGCGTCAGGGGGACGAG	3405
1137	QY	CysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyr	1156
3406	DB	TGCCAGCTATGTGAGGTAGAAAAATCGATACCAAGGAAACCTCTCTCAGAGGAACATGTTAT	3465
1157	QY	TyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyr	1176
3466	DB	TATACTCTTCTTATTGTACTATCAGTTTCACTTTAGTCTATCCACGAGAGATGATCGCTAT	3525
1177	QY	TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe	1196
3526	DB	TACACAGCTATCAATTTTGTGGGTACTCTCTGACGACAAACACAGGATTTGGACATGTTTC	3585
1197	QY	IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGly	1216
3586	DB	ATCAATGCCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTTCTCAGCTGGA	3645
1217	QY	ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAsp	1236
3646	DB	ACCCAGGCTGCAGAAAGATGCCTGTGTGTTTCAAAACCAACATTAAGGAGTACAAAGAT	3705
1237	QY	SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr	1256
3706	DB	AGTTTCTCTAATGAGAGTTTGATTTTCGCAACCAACCAATATCATCTTCTTTGTTTAT	3765
1257	QY	ValSerAsnPheThrTrpProfileLysIleGlnIleAlaPheSerGln	1272
3766	DB	GTCAAGTAAATTTCACTTGGCCCATCAAAATTCAGGTGCAAACTCAACAA	3813

## RESULT 4

US-10-197-824-1	Sequence 1, Application US/10197824	DB
US-10-197-824-1	Sequence 1, Application US/10197824	982
US-10-197-824-1	Publication No. US20040023219A1	QY
US-10-197-824-1	GENERAL INFORMATION:	
US-10-197-824-1	APPLICANT: ORIGENE TECHNOLOGIES INC	DB
US-10-197-824-1	TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES	
US-10-197-824-1	FILE REFERENCE: NOVEL PROSTATE CANCER GENES	
US-10-197-824-1	FILE REFERENCE: 9U 106 R1	QY
US-10-197-824-1	CURRENT APPLICATION NUMBER: US/10/197,824	DB
US-10-197-824-1	CURRENT FILING DATE: 2002-07-19	
US-10-197-824-1	NUMBER OF SEQ ID NOS: 39	QY
US-10-197-824-1	SOFTWARE: Patent in version 3.1	
US-10-197-824-1	SEQ ID NO 1	DB
US-10-197-824-1	LENGTH: 6733	
US-10-197-824-1	TYPE: DNA	
US-10-197-824-1	ORGANISM: Homo sapiens	QY
US-10-197-824-1	FEATURE:	
US-10-197-824-1	NAME/KEY: CDS	DB
US-10-197-824-1	LOCATION: (575)..(4711)	
US-10-197-824-1	OTHER INFORMATION:	QY
US-10-197-824-1		

Db 1282 TAAATACGGAAGGGTGAAGCTTGTGATATTCCTTACTGTAAGCCAATTGGCGAGTCC 1341  
Qy 294 phi.sarggly.leCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpG1 314  
Db 1342 AGATCAGCGTTACTGTGACCTGACCTGAGGAGAAAATATGTGTCTGCAATGATAGTGGCA 1401  
Qy 314 nGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTy 334  
Db 1402 AGGTCCTGATGTCTTTGAATGTTCCCTCTACTGAGCTTACTGGATT-----CT 1452  
Qy 334 sSerAsnLeuLys-----LeuProArgAlaSerHisIleValValValAs 349  
Db 1453 GCCAACGTTAAACCCCTTCAGTCCTCTGTAGGTCGGCTTCACATAAGCAGTTTACCA 1512  
Qy 349 nGlyAsnIleMetTrpValValGlyGlyTyMetPheAsnHisSerAspTrpAsnMetVa 369  
Db 1513 CGGGAATTTAATGCGGGTGATGTGGGATATCTTTTAACTACAGTCTTTTCAATGGT 1572  
Qy 369 lleuAlaTyRAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnVa 389  
Db 1573 CTTAAATACAAATTTAGAAAGCAGTATATGG---AATGTAGGAATCCCATCAAGGGACC 1629  
Qy 389 lValValArgTyGlyHisSerLeuAlaLeuTyRlysAspLysIleTyMetTyRGlYl 409  
Db 1630 TCTCAGAGATATGACACTCTCTTCTTTATATCAGGAACAACATCTTTATGTATGGAG 1689  
Qy 409 llysIleAspProThr---GlyAsnValThrAsnGluLeuArgValPheHisIleHisAs 428  
Db 1690 CAGAATTCAAACAAATGATGGCAATGTCCACAGATGAATATATGGTTTTAAACATACATAG 1749  
Qy 428 nGluSerTrpValLeuLeuThrProLysAla-----LysGluGlnTyRAlaValVa 445  
Db 1750 TCAGTCATGGAGTACAAAACCTCTACTGTCTTGGACATGGTCAGCATGTCTGTGGA 1809  
Qy 445 lGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePh 465  
Db 1810 GGGCAATTCAGCACATATATGAGAGTGGATAGATGATGTGTGATGATCAATAATTT 1869  
Qy 465 eGlyHisCysProLeuTyRGlYTyRlSerAsnValGlnGluTyRAspLeuAspLysAs 485  
Db 1870 TGGATATCTGCAATATATGTTATATCAAGCAGCATACAGGAATACCATATCTCTCAAA 1929  
Qy 485 nThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyRGlYHisSerSe 505  
Db 1930 CACTTGGCTGTTCAGAACTAAGAGAGCTATTGTACAGGTGATATGCCATATAG 1989  
Qy 505 rValTyRAspHisArgThrArgAlaLeuTyRValHisGlyTyRlYsAlaPheSerAl 525  
Db 1990 TGTGTATGATGAATAAACAAGTCCATTTATGTTTCATGGGGTATAAAGCATTGCCAGG 2049  
Qy 525 aAsnLysTyRArgLeuAlaAspAspLeuTyRArgTyRAspValAspThrGlnMetTrpTh 545  
Db 2050 GAACAAATATGGATTGGTTGATGATCTTTATAAATATGAATTAACACATAAGACTTGGAC 2109  
Qy 545 rIleLeuLysAspSerA-gPhePheArgTyRleuHisThrAlaValIleValSerGlyTh 565  
Db 2110 TATTTTGAAGAAGAGTGGTTTGGCAGATACCTTCATTCAGCTGTCTTCAATGAGNC 2169  
Qy 565 rMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCy 585  
Db 2170 TATGCTTATTTTGGAGGAAATACCCATAATGACACTTCCTTTAGTAGTACCGTGCAAAATG 2229  
Qy 585 sPheSerSerAspPheMetAlaTyRAspIleAlaCysAspArgTrpSerValLeuProAr 605  
Db 2230 TTTTCTCCGGAATTTCTGGCATATGACATAGCTTGTGTGATGAATGGAAATATCTACCAA 2289  
Qy 605 gProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerTh 625  
Db 2290 ACCAAATCTTCATAGAGATGTCAACAGATTTGGACACTCTCAGTAGTCAATTAACGGGTC 2349  
Qy 625 rMetTyRValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSe 645  
Db 2350 CATGTATATATATTTGGGGGATTTCTTAGTGTACTCTTAATGATATCTCTTGTATACAAGCC 2409

Qy 645 rGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleAr 665  
Db 2410 TCCAAATTCGACGCTTTTCAGAGATGAAGAACTTTGTAAAAATGCTGGTCCAGGATAA 2469  
Qy 665 sCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluG1 685  
Db 2470 ATGTGTTTGGAAAT-----AAAAATCAGTGTGAATCTTGG-----GAATCTGGGAATC 2517  
Qy 685 nGluGluLeuLysSerGluCysPheSerLysArgThrLeuAspHisAsnArgCysAs 705  
Db 2518 TAATAATATTTCTTAGAGCAAAAGTCCCTCTTAAACAGCTGCTTCTGATGACAGATGTTA 2577  
Qy 705 pGlnHisThrAspCysTyR-SerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAs 725  
Db 2578 CAGATATGCAAGTTGTGCCAGCTGTACTGCCAATACAATGGGTGCCAATGGTGTGATGA 2637  
Qy 725 p---HisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTy 744  
Db 2638 CAAGAAATGCAATTCGGCAAAATAGTAACAGT-----ATGCTGTCAAGAACTA 2688  
Qy 744 rGluAsnCysProLysAspAsnProMetTyRtyrCysAsnLysLysThrSerCysArgSe 764  
Db 2689 CACCAATATGTCATGTGAGAAATGAGCAG--ATTGTAAACAACCTTACAGCTGTAAAAG 2745  
Qy 764 rCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPr 784  
Db 2746 CTGTTCACATAAATCTGAATTCAGTGGATCAGACAGACAGCAAGATGCCAGCTTTTACC 2805  
Qy 784 oGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrTh 804  
Db 2806 AGTCACTCTTTTGGGAAGAGTGGAGTCATATTTGGGATGCTTGTCTTAGTCAATTC 2865  
Qy 804 rAlaLysGluAsnTyRAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAl 824  
Db 2866 CACTAGAGAAATCATGACATGCAAACTTTATTCGTATATCTTAGTGGAAATCTTGC 2925  
Qy 824 aSerLeuThrThrGlnLysValGluPheValLeuLysGlnLysGlnLeuArgIleMetGlnSe 844  
Db 2926 TTCATTAAACAACCTCAAAAAGTAGATATTTGTTCTGGATGAAATA-----2971  
Qy 844 rSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnVa 864  
Db 2972 ----CAGAAGTATACACACAGAAAGTATCACCTTGGTAGGCTTGGCAGAGTCAATAT 3027  
Qy 864 lSerTyRTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPr 884  
Db 3028 ATCTTATTTGGGATGGGAAGACATGCTCTCTTTTACAAACACACACTACAGTGGCTTCC 3087  
Qy 884 oSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLe 904  
Db 3088 TGGCGAACCCCAATGATTTCTGGGTTTTTGTGCATATCTGGAAGGGCTGCGAGTGGCAGCTT 3147  
Qy 904 uLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAla-----922  
Db 3148 AAAAGCTAATCTTGTATACATCTATGCAAAATGGCCCTTGTCTGTGAAAAACCTGTTTAG 3207  
Qy 923 ----AsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAs 941  
Db 3208 TCCAAATACAAATGGAGGCCCTGCAAAAAGCCATGCTCTCTGAGGACATCATGTTCCAA 3267  
Qy 941 pCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSe 961  
Db 3268 CTGTCACGAATGGCATGGAGTGTATGTGTGTCAGCAGTACGAAACGATGTTGTGACTC 3327  
Qy 961 rAsnAlaTyRValAlaSerPheProPheGlyGlnCysMetGluTrpTyRThrMetSerTh 981  
Db 3328 TAATGCTATATCATCTCTTTCCATATGGACAAATGTCTAGAGTGGCAAACT---GCCAC 3384  
Qy 981 rCysProGluAsnCysSerGlyTyRtyrCysThrCysSerHisCysLeuGluLeuProGl 1001  
Db 3385 CTGCTCCCTCTMAATTTGTTCTGGATGAGAACCCTGTGGACAGTGTITGGAAACAGCCTGG 3444

1001	Y	Cys	Gly	T	Trp	Cys	Thr	Asp	Pro	Ser	Asn	Thr	Gly	Ser	Gly	Ser	Leu	Gly	Ser	Ty	1021
3445	A	T	G	G	C	T	G	C	T	A	G	T	A	T	A	C	A	G	A	G	3504
1021	r	Lys	Gly	P	ro	Val	Lys	Met	Pro	Ser	Gln	Ala	Pro	Thr	Gly	Asn	Phe	Tyr	Pro	Gln	1041
3505	A	C	G	G	G	A	C	C	A	A	T	G	A	G	C	T	T	-----	A	T	3549
1041	u	Leu	Asn	Ser	Ser	Met	Cys	Leu	Glu	Asp	Ser	Arg	Tyr	Asn	Trp	Ser	Phe	Ile	His	Cys	1061
3550	T	C	T	T	G	A	C	C	A	A	T	T	T	G	C	C	C	A	A	A	3609
1061	o	Ala	Cys	Gln	Cys	Asn	Gly	His	Ser	Lys	Cys	Ile	Asn	Gln	Ser	Ile	Cys	Glu	Lys	Cys	1081
3610	A	G	T	T	C	C	C	A	G	T	A	A	T	G	C	A	T	A	A	T	3669
1081	u	Asn	Leu	Thr	Thr	Gly	Lys	His	Cys	Glu	Thr	Cys	Ile	Ser	Gly	Phe	Tyr	Gly	Asp	Pro	1101
3670	A	A	A	T	C	T	C	A	C	A	G	A	A	G	C	A	G	T	G	T	3729
1101	r	Asn	Gly	Gly	Lys	Cys	Gln	Pro	Cys	Lys	Cys	Asn	Gly	His	Ala	Ser	Leu	Cys	Asn	Thr	1121
3730	C	A	A	T	G	T	G	G	A	C	A	G	T	T	G	A	C	A	A	T	3789
1121	n	Thr	Gly	Lys	Cys	Phe	Cys	Thr	Thr	Lys	Gly	Val	Lys	Gly	Asp	Glu	Cys	Gln	Leu	Cys	1141
3790	C	A	C	A	G	A	A	A	T	G	T	T	C	G	C	A	C	T	A	A	3849
1141	u	Val	Glu	Asn	Arg	Tyr	Gln	Gly	Asn	Pro	Leu	Arg	Gly	Thr	Cys	Tyr	Tyr	Thr	Leu	Leu	1161
3850	C	T	C	T	G	A	A	A	T	C	G	T	A	T	T	T	T	T	T	T	3909
1161	e	Asp	Tyr	Gln	Phe	Thr	Phe	Ser	Leu	Ser	Gln	Glu	Asp	Asp	Arg	Tyr	Tyr	Thr	Ala	Ile	1181
3910	T	G	A	A	T	A	C	A	A	T	T	A	C	A	G	A	G	A	T	G	3969
1181	n	Phe	Val	Ala	Thr	Pro	Asp	Glu	Gln	Asn	Arg	Asp	Leu	Asp	Met	Phe	Ile	Asn	Ala	Ser	1201
3970	C	T	T	T	A	T	A	G	A	A	A	A	A	A	A	T	C	G	A	T	4029
1201	s	Asn	Phe	Asn	Leu	Asn	Ile	Thr	Trp	Ala	Ile	Ser	Phe	Ser	Ala	Gly	Thr	Gln	Ala	Gly	1221
4030	C	A	A	C	T	T	T	A	A	C	T	T	C	T	G	T	G	T	C	A	4089
1221	u	Glu	Met	Pro	Val	Ile	Ser	Lys	Thr	Asn	Ile	Lys	Glu	Tyr	Lys	Asp	Ser	Phe	Ser	1241	
4090	A	G	A	C	A	T	T	C	T	A	G	A	A	T	A	A	A	G	A	A	4149
1241	u	Lys	Phe	Asp	Phe	Arg	Asn	His	Pro	Asn	Ile	Thr	Phe	Phe	Val	Tyr	Val	Ser	Asn	1261	
4150	A	A	A	A	T	T	A	A	C	T	T	A	G	A	C	A	T	T	C	A	4209
1261	r	Trp	Pro	Ile	Lys	Ile	Gln	Ile	Ala	Phe	Ser	Gln	His	Ser	Asn	Phe	Met	Asp	Leu	1281	
4210	C	T	G	C	C	T	T	A	A	A	T	C	T	C	A	C	A	C	A	A	4269
1281	n	Phe	Phe	Val	Thr	Phe	Phe	Ser	Cys	Phe	Leu	Ser	Leu	Leu	Val	Ala	Ala	Val	Val	1301	
4270	G	T	T	T	T	T	G	T	C	A	G	T	T	T	T	T	T	T	T	T	4329
1301	p	Lys	Ile	Lys	Gln	Ser	Cys	Trp	Ala	Ser	Arg	Arg	Glu	Gln	Leu						

```

Db      4510  TGTGGGAACAGAGCTGCTTCTTGACTGTGTTTCTTTCTTACACGAGGATCATCAGG 4569
Qy      1381  yileProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGI 1401
Db      4570  TGGCCCTCCCTCCCTGGGCAGTCAGGCCCTTGCATTTGCCAGTGCCTAATAGATATTTTACA 4629
Qy      1401  nGlnMetProIleValTyrIysGlnIySerGlyAlaValArgAsnArgIysGlnGlnPr 1421
Db      4630  ACAGAAAGCTTCAGATAGTAAGATTAAGACTTCTGAGTGCCGGAATCGAAACACCTTTC 4689
Qy      1421  oProAlaGlnProGlyThrCysIle 1429
Db      4690  AACACGCTCAA--GGAACCTTGTC 4711

RESULT 5
US-09-893-238-18
; Sequence 18, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 18
; LENGTH: 2625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-893-238-18

Alignment Scores:
Pred. No.: 0 Length: 2625
Score: 4136.50 Matches: 759
Percent Similarity: 95.48% Conservative: 2
Best Local Similarity: 95.23% Mismatches: 22
Query Match: 52.26% Indels: 14
DB: 9 Gaps: 3

US-09-787-097-12 (1-1429) x US-09-893-238-18 (1-2625)
Qy      2 ValAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAlaAla 21
Db      10 GTGGCCGACGCGCGGCACTGAGGCAAGGCTCAGGAGGAGGACGGCGGCGAGCGGACGC 69
Qy      22 LeuAlaGlyArgSerGlyGlyProHisTirPspTirPspValThrArgAlaGlyArgPro 41
Db      70 CTGGGGGCGAGGAGCGGCGGCGCGCCGACCCGACCCGACGAGGGGCTTGGAGGCGG 129
Qy      42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56
Db      130 GGAACCGCGCGCGCGGTGTGTCTCCGCGGGGTGCTGTCGGGGGCGGTGCCCGCGCGCG 189
Qy      57 LeuArgProArgLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
Db      190 CTGCTGCGG---CTGCTCTTTTTCGCTGCTGCTGCGG----- 225
Qy      77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96
Db      226 CTGCCCCGGGAGGCGGAGGCGGTGGCGGTGGCGCGCGCGGTGTCGCGCTCGGCGCGAC 285
Qy      97 GluAlaIyGlnCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116

```

286	Db	GAGGCCAAGGAATGTGACCGCGCGTGTGTCAACGGCGGTGCTGCTGCAACCCCTGGCACCGCGC	345
117	Qy	GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe	136
346	Db	CAGTGCGTCTCCCGCGCGCGTGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGTTC	405
137	Qy	ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr	156
406	Db	AGACTAATCTGGATCTCTCGGGTGTGGACAGATGGACCTGGAAATTTAAATACAAAACG	465
157	Qy	LysCysThrTrpLeuLleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis	176
466	Db	AAGTGCAGCTGCTCATTTGAAGGACAGCAAAATAGAATAATGAGACTTCGTTTCAATCAT	525
177	Qy	PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla	196
526	Db	TTTGCTACAGAGTGTAGTGGACCATTTATGTGTTTATGATGGGACCTCAATTTATGCA	585
197	Qy	ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal	216
586	Db	CCGTAGTGTGTGATTTAGTGGCTCATTTGTTCTGAGAGAGATGGCAATGAGACTGTC	645
217	Qy	ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr	236
646	Db	CCTGAGGTTGTGGCACATCAGTTATGCTTCATTTTGTAGTATGCTGCTTAT	705
237	Qy	AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly	256
706	Db	AATTGACTGGATTTAATTACTTTACAGTTTGTATGTGTCCAAATAACTGCTCAGGC	765
257	Qy	ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn	276
766	Db	CGAGGAGTGTAAAGATCAGTAATAGCAGCGATACTGTTGAATGTGAATGTTCTGAAAC	825
277	Qy	TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg	296
826	Db	TGGAAAGGTGAAGCATGTGACATTCCTACCTGTACAGCAACTGTGGTTTCTCATCGA	885
297	Qy	GlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro	316
886	Db	GGCATCTGCAATTCAGTGATGTCAGAGGATGCTCTCTGCTCTCAGACTGGCAGGCTCT	945
317	Qy	GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn	336
946	Db	GGATGTTTCAGTCTCTACCAAGCTAAACCAAGTCATTTGGACTCGAGAGGAATATCTTAAC	1005
337	Qy	LeuLysLeuProArgAlaSerHisIleValValValAsnGlyAsnIleMetTrpValVal	356
1006	Db	TTAAAGCTCCCGAGAGCATCTCAATAAGCTGTGTCAATGGAAACATTTATGTGGGTGTT	1065
357	Qy	GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer	376
1066	Db	GGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCTAGCGTATGACCTTGCTTCT	1125
377	Qy	ArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer	396
1126	Db	AGGAGTGGCTTCACTAAACCGTCTGTGAACAAATGTGGTTGTTAGATATGGTCATTCT	1185
397	Qy	LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn	416
1186	Db	TTGGCATTTATACAGGATATAATTTACATGTATGGAGGAAAAATTTGATTCACCTGGGAAT	1245
417	Qy	ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrPro	436
1246	Db	GTGACCAATGATGTAGAGTTTTTCACATTCATTAATGAGTCATGGGTGTTGTTGACCCCT	1305
437	Qy	LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn	456
1306	Db	AAGCAAAAGGAGCAGTAGTCAGTGGTGGGCACCTCTGCACACATTTGTACACTGAAGAAT	1365
457	Qy	GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn	476
1366	Db	GGCCGAGTGGTCATGCTGGTTCATCTTTGGTCACCTGCTCTATGATATATGAACAAT	1425

Qy	477	ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu	496
Db	1426	GTGCAGGAATATGATTTGGTAAGAACACATGAGGATATATTACACACCCAGGGTGCCCTT	1485
Qy	497	ValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal	516
Db	1486	GTGCNAGGGGTACCGCCATAGCAGTGTTTACGACONTAGGACCGGGCCCTATACGTT	1545
Qy	517	HisGlyGlyTyrIysAlaPheSerAlaAsnLysTyrArgLeuAlaAspLeuTyrArg	536
Db	1546	CATGGTGGCTACAAAGGCTTTTCAGTGGCCATAAAGTACCGGCTTCAGATGATCTCTACCGA	1605
Qy	537	TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPheArgTyrLeu	556
Db	1606	TATGATGTGATACCCAGATGTCACATCTTTAAGACAGCCGATTTTTCCGTTACTTG	1665
Qy	557	HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp	576
Db	1666	CACACAGCTGTGATAGTGAGTGGAAACCATCTGGTGTTCGGGGAAACACACACAATGAC	1725
Qy	577	ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla	596
Db	1726	ACATCTATGAGCCATGGCGCAATGCTTCTCTTCAGATTTTCATGGCTATGACATTGCC	1785
Qy	597	CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly	616
Db	1786	TGTGACCGCTGGTCAGTGCTTCCAGACCTGATCTCCACCATGATGTCAACAGATTGGC	1845
Qy	617	HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu	636
Db	1846	CATTACAGATCTTACACACAGACACCATGATGTGTTCGGTGGTTTCAATATGTCCTC	1905
Qy	637	LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla	656
Db	1906	CTCAGCGACATCCTGGTATTACCTCGGAACAGTGTGATCGCATCGAGTGAAGCGCT	1965
Qy	657	CysLeuAlaIalaglyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle	676
Db	1966	TGTTTACACAGCAGACCTGCTATTTCGGTGTGTGTGGAACACAGGGTCGTCTCAGTGTATC	2025
Qy	677	SerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys	696
Db	2026	TCGTGGCGCTGGCNACTGATGACACAGAGAAAGTTAAATCAGAAATGTTTTCCAA	2085
Qy	697	ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn	716
Db	2086	AGAACTCTTGACCATGACAGATGTGACCAGCACACAGATTGTTACAGCTGCACAGCAAC	2145
Qy	717	ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu	736
Db	2146	ACCATGACATGCCACTGGTGCAATGACCATTTGTGTCCCGAGAACACACAGCTGCTCAGAA	2205
Qy	737	GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys	756
Db	2206	GGCCAGATCTCCCATTTTATGATATGAAATGGCCCAAGGATTAACCTATGTACTACTGT	2265
Qy	757	AsnLysIleThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg	776
Db	2266	AACAAGAACACAGCTGCAGAGCTGTGCCCTGGACACAGAACTGCCAGTGGGAGCCCCGG	2325
Qy	777	AsnGlnGluCysIleAlaLeuProGluAsnIleCys-GlyIleGlyTrp	792
Db	2326	AATCAGAGTGCATTCCTTCGCCGGTGGGCCCTTTCAGGGCTCATCTTGG	2374

RESULT 6  
US-09-893-238-12  
; Sequence 12, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: MOORE, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
 ; FILE REFERENCE: 7853-237  
 ; CURRENT APPLICATION NUMBER: US/09/893,238  
 ; CURRENT FILING DATE: 2001-06-27  
 ; PRIOR APPLICATION NUMBER: 09/245,041  
 ; PRIOR FILING DATE: 1999-02-05  
 ; PRIOR APPLICATION NUMBER: 60/093,630  
 ; PRIOR FILING DATE: 1998-07-21  
 ; PRIOR APPLICATION NUMBER: 60/104,978  
 ; PRIOR FILING DATE: 1998-10-20  
 ; NUMBER OF SEQ ID NOS: 129  
 ; SOFTWARE: FastSeq For Windows Version 3.0  
 ; SEQ ID NO 12  
 ; LENGTH: 6370  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-09-893-238-12

Alignment Scores:  
 Pred. No.: 0 Length: 6370  
 Score: 3491.50 Matches: 656  
 Percent Similarity: 95.51% Conservative: 3  
 Best Local Similarity: 95.07% Mismatches: 18  
 Query Match: 44.11% Indels: 13  
 Gaps: 4

US-09-787-097-12 (1-1429) x US-09-893-238-12 (1-6370)

QY 753 MetTyrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGln 772  
 Db 1 ATGTACTACTGTAAACAAGACACCAAGCTGCAGAGCTGTGCCCTGGACCAAGCTGCCAG 60  
 QY 773 TrpGluProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrp 792  
 Db 61 TGGGAGCCCGGAAATCAGAGGTGCATGCCCTGCCGGAATAATCTGTGGCATTGGCTGG 120  
 QY 793 HisLeuValGlyAsnSerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAla 812  
 Db 121 CATTGGTTGGAACTCATGTTTGAATTAATCTACTGCGAGGAGAAATATGACATGCT 180  
 QY 813 LysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrGlnLysLysVal 832  
 Db 181 AAATTGTTCTGTAGGAACCAACATGCCCTTTTGGCTTCTTTACAAACCCAGAGAAGGTA 240  
 QY 833 GluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThr 852  
 Db 241 GAATTTGTCCTTAAGCAGCTGCGAATAATGCAAGTCAATCTCAGAGCATGTCCAGCTCACC 300  
 QY 853 LeuThrProTrpValGly---LeuArgLysIleAsnValSerTyr-TrpCysTrp-GluA 871  
 Db 301 TTAACCCCATGGTTCGGGCTTCGGGAAGGTCAATGTGCTACTKGGTGTGGGGAAG 360  
 QY 871 spMetSerPro---PheThrAsnSer-LeuLeuGlnTrpMetProSerGluProSerAsp 889  
 Db 361 GATATGKTCCTCCATTTTACAAATAGTTTATTACAGTGGGATGSCCGCTTCGAGCCCGAG 420  
 QY 890 Ala-----GlyPheCysGly---IleLeuSerGluProSer---ThrArgGlyLeuLysA 906  
 Db 421 TGTGCTTGGRATCTGTGGGAATTTATTTCAGGAACCCAGTTACTTCGGGAGCTGAAGG 480  
 QY 906 lAlaThrCysIle-AsnProLeu-AsnGlySerValCysGluArgProAlaAsnHisSe 925  
 Db 481 CTGCAACCTGCAATCAACCACTTAAATGTGTAGTGTGTGTGAAGCCCTGCAAAACCAAG 540  
 QY 925 rAlaLys-GlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerG 945  
 Db 541 TGCATAGGCAGTGCCTGGACACCACTGTGCTTGAGGACGAGCATGTGAGATTCCACACCG 600  
 QY 945 LysSerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTrpV 965  
 Db 601 GCAGCTCTCAGTGCATGTGGTGCAGCAACATGAAGCAGTGTGTGACTCCAATGCTATG 660  
 QY 965 aAlaSerPheProPheGlyGlnCysMetGluTrpTyr-ThrMetSerThrCysProProG 985

Db 661 TGGCCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGACCTCCCCCCTG 720  
 QY 985 luAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyTrpC 1005  
 Db 721 AAAATTGTTTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGGT 780  
 QY 1005 ystThrAspProSerAsnThrGlyLysGlyCysIleGluGlySerTyrLysGlyProV 1025  
 Db 781 GTACTGATCCAGCAATACTGCAAAAGGAAATGCAATAGAGGTTCTCTATAAAGGACAG 840  
 QY 1025 allysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerS 1045  
 Db 841 TGAAGATGCTTCGCAAGCCCTCAGAGAAATTTCTATCCACAGCCCTGCTCAATTCCA 900  
 QY 1045 erMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnC 1065  
 Db 901 GCATGTGCTAGAGGACAGACATACAACTGCTCTTTCAATTCACGTCCAGCTTCCCAAT 960  
 QY 1065 ysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrT 1085  
 Db 961 GCAACGCGCACATAAATGCAATCATCAGAGCATCTGTGAGAGTGTGAGAACCTTGACCA 1020  
 QY 1085 hrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyL 1105  
 Db 1021 CAGGCAAGCAGCTGCAGACCTGCATATCTGGCTTCTACGGTGTATCCCAATGGAGGA 1080  
 QY 1105 yscysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysC 1125  
 Db 1081 AATGTGAGCATGCAAGTGCATGGCAGCGCTCTGTGCAACCAACACAGGGCAAGT 1140  
 QY 1125 ysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnA 1145  
 Db 1141 GCTTCTGCACCAAGGCGGTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAAAATC 1200  
 QY 1145 rgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnP 1165  
 Db 1201 GATACCAAGAAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTATTGTACTCATG 1260  
 QY 1165 heThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheValAlaI 1185  
 Db 1261 TCACCTTTAGTCTATCCAGGAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTA 1320  
 QY 1185 hrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnL 1205  
 Db 1321 CTCCTGACGAACAAACAGGAGTTTGACATGTTCAATCAATGCTCCAAAGAAATTTCAACC 1380  
 QY 1205 euAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProV 1225  
 Db 1381 TCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGCCCTG 1440  
 QY 1225 alValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspP 1245  
 Db 1441 TTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAATTTGATT 1500  
 QY 1245 heArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleL 1265  
 Db 1501 TTCGCAACCAACCAATATACATCTTTGTTTATGTCAGTAATTTTACCTGCCCCATCA 1560  
 QY 1265 ysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValT 1285  
 Db 1561 AAATTTCAGATTGGCTTCTCTCAGCACAGCAATTTTATGACCTGTGTACAGTTCTTCGTGA 1620  
 QY 1285 hrPhePheSerCysPheLeuSerLeuLeuValAlaIleValTrpLysIleLysG 1305  
 Db 1621 CTTTCTTCACTTGTCTCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1680  
 QY 1305 lnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaL 1325  
 Db 1681 AAAGTTGTTGGGCTCTCAGACGTGAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCA 1740  
 QY 1325 erArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuI 1345

Db 1741 GCCGTCCTTTGCTCTGTAATGTCGCTTGGAAACAGATGAGGAGCCTCTGATCTTA 1800  
Qy 1345 leglyGlySerlleysThrValProlysProIleAlaLeuGluProCysPheGlyAsnL 1365  
Db 1801 TTGGGGGAGATAAAGACTGTTCCTCCAAACCATTCACACTGAGCCGTGTTTTGGCAACA 1860  
Qy 1365 ysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyLeuGlyProProp 1385  
Db 1861 AAGCGCTGTCTCTCTGCTGTTGTGAGGCTCCCTCGAGGCTGGTGGCATCCCTCTC 1920  
Qy 1385 roGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProI 1405  
Db 1921 CTGGCAGTCAGTCTGCTGTGCCACGCGCTGGTGGACATTCACAGATGCCGA 1980  
Qy 1405 leValTylsGlyLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnP 1425  
Db 1981 TAGTGTACAAGAGAGAGTCAGGAGCCGTGAGAACCGGAGCAGCAGCCCTGCACAGC 2040  
Qy 1425 roGlyThrCysIle 1429  
Db 2041 CTGGACCTGCATC 2054

## RESULT 7

US-09-893-238-8  
; Sequence 8, Application US/098933238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-8

## Alignment Scores:

Pred. No.:	0	Length:	2419
Score:	2435.00	Matches:	642
Percent Similarity:	84.37%	Conservative:	22
Best Local Similarity:	81.58%	Mismatches:	44
Query Match:	43.40%	Indels:	79
DB:	9	Gaps:	4

US-09-787-097-12 (1-1429) x US-09-893-238-8 (1-2419)

Qy 2 VallAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAlaAla 21  
Db 83 GTGGCGGGCGGGCGGCGACTGAGCGCGGTGAGGGGGAGCAGGAGCAGCAGCAGCG 142  
Qy 22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41  
Db 143 CTGGCGGCGAGAGGCGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 202  
Qy 42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56  
Db 203 GGACCGCGCGCGCGCTGTGTCTCCGCGGGTGTGTGCGGGCGCTGCGCGCGCGCGCG 262  
Qy 57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeu 76  
Db 263 CTGTGCGCG---CTGCTCTTTTGTGCTGCTGCTGCTGCGG----- 298

Qy 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96  
Db 299 CTCCCCCGGAGCGCCAGGCCGCTGCGGTGGCGCGCGCGGTGTCGGCTTCGGCCGAGCC 358  
Qy 97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116  
Db 359 GAGGCCAAGGAATGTGACCGCGCGTGTGTCAACGGCGCGCGCTGCAACCTCGCACCGGC 418  
Qy 117 GlnCysValCysProAlaGlyTrpValGlnGluGlnCysGlnHisCysGlyGlyArgPhe 136  
Db 419 CAGTGGCTCTGCCCCACCGGGCTGGGTGGCGAGCAATGCGACACTGCGGGGGCGCTTC 478  
Qy 136 ----- 136  
Db 479 AGGACATCTGTCTCACGCTATAATCACAGCTGTTTCGGAAGGTGAGGCTGGAGGAACACT 538  
Qy 136 ----- 136  
Db 539 TCGAGGCAAGCTTCGGCTACAGAATAAGTTCAAGAGTAACCTGGGGCAACTTGGGCTTGT 598  
Qy 136 ----- 136  
Db 599 CTCCAAACCAAAATGACGCAAAAGGAGCAAGCTAGACTCTTTTCGGAAAAATTTAGCTG 658  
Qy 137 -----Arg-LeuThrGlySerSerGlyPheValThrAspGlyProGlyAs 151  
Db 659 ACTAATTTTTCACCGAGAACTAATCGCTCTTCTGGAATTTGTAACAGATGACCTGGGAA 718  
Qy 151 nTylsTylsTylsThrLysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetAr 171  
Db 719 TTATAAATATAGACGAAGTGACATGGCTCAATTAAGSACAGCCAAATAGAAATATAG 778  
Qy 171 gLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspG 191  
Db 779 ACTTCGCTTCAACCAITTTTGTACAGATGTAGCTGGGACCAATTTATATGTTATGATGG 838  
Qy 191 YaspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAs 211  
Db 839 GGACTCAATCTACGCACCTCTGATTCGCTCTTATGTCGCTCATTTGTTCCCTGGAAGAGA 898  
Qy 211 pGlyAsnGluThrValProGluValAlaAlaThrSerGlyTyrAlaLeuLeuHisPhePh 231  
Db 899 TGGCATGAGACGGCTCTGAGGTCACTGCTCACTTCAGGTTATGCACTGCTGCAATTTT 958  
Qy 231 eSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysPr 251  
Db 959 CAGTGATGCTGCTTATAATCTGACTGGATTTAATATCACTTACATTTTGACATGTGTC 1018  
Qy 251 oAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCy 271  
Db 1019 GAATAATGCTCAGCGCGAGGAGAGTGAAGAGCAGTACACAGCAGCGCTGTTGAGTG 1078  
Qy 271 sGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCy 291  
Db 1079 TGAATGTTCTGAAAACCTGGAAGGGGAGTGTGTGACATTCCTCACTGTACAGACAACGT 1138  
Qy 291 sGlyPheProHisArgGlyLysCysAsnSerSerAspValArgGlyCysSerCysPheSe 311  
Db 1139 TGGCTTCTCTACCGAGGCACTCTGTAATGCAAGCAGCATACAGAGGGTGTCTCTGTTTC 1198  
Qy 311 rAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrAr 331  
Db 1199 TCACTGGCAGGGTCTCGGATGTTCAATTCCTGTGCCAGCTAACCACTCTTTTGGACTCG 1258  
Qy 331 gGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAs 351  
Db 1259 AGAAGAATATCTGATTTAAAGCTTCCCAAGAGCCTCTCAATAAGCTGTGTCAATGGA 1318  
Qy 351 nIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAl 371  
Db 1319 TATAATGCGGTGTTGGCGGATATATGTTCAACCAATTCAGATTACAGCATGCTTTTACG 1378

```
QY 371 aTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValVal 391
Db 1379 GTATGACCTGACTTCTAGGAATGCTTCCATTAACCATTTCTGTGAACAGTGGTGTG 1438
QY 391 lArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysI 411
Db 1439 AAGATATGCTCAATCTTTGGCAATTAACAAGATTAATATCTACATGTATGGAGGAAAT 1498
QY 411 eAspProThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTr 431
Db 1499 TGATTCACAGGAGGACGAGACCAATGAGCTGAGAGTATTTTCATATTCATATGAATCATG 1558
QY 431 pValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisI 451
Db 1559 GGTATTTGTTAACTCCGAAAGCTAAGATCAGTATGAGTGTGTGACACTCAGACACAT 1618
QY 451 eValThrLeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTy 471
Db 1619 TGTTACACTGGCATCTGGCCGTGTGTGATGTGTGTCATCTTCGTCATTCCTCCACTCTA 1678
QY 471 rGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHi 491
Db 1679 TCGATATATAGCGTTGTGCAGAAATATGACTTGGAAAGAACACATCGAGTATATTACA 1738
QY 491 sThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerValTyrAspHisArgTh 511
Db 1739 TACTCAGGCTGCTCTGTGTCAAGGGGGTGTATGCCACAGTAGTGTTTATGATGACAGGAC 1798
QY 511 rArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAl 531
Db 1799 CAAGCTCTGTAGCTTCATGGTGGCTACAGGCTTTCAGCGCCCAACAAATACCGGCTTC 1858
QY 531 aAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerAr 551
Db 1859 AGATGACCTCTACAGATACGATGTGCATACCTAGATGTGGACCATTTCTTAAGACAGCG 1918
QY 551 gPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGl 571
Db 1919 ATTTTTCGTTACTTGCATACAGCTGTGTATAGTGTAGTGGACCATGCTGGTGTGTGGAGG 1978
QY 571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591
Db 1979 GAACACACAAATGACACTTCCATGAGCCACCGTCCCAATGCTTCTCTCCGACTTCAT 2038
QY 591 tAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisHisAs 611
Db 2039 GGCCTTATGACATGCTTGTGACCGATGTGTAGTGTAGTGGACCATGCTGGTGTGTGGAGG 2098
QY 611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGl 631
Db 2099 TGTCAACAGATTTGGCAATTCAGCAGTCTGTATACACAGCAGCAGCATGTATGTTCGGCGG 2158
QY 631 yPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHi 651
Db 2159 CTTCAACAGCTTCTCTCCTCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2218
QY 651 sArgSerGluAlaAlaCysLeuAlaGlyProGlyIleArgCysValTrpAsnThrGl 671
Db 2219 CCGCAGTGAAGCTGCTGTGTGGCAGCAGGACCTGGTATCGGTGTCTGTGTGGACACACA 2278
QY 671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysLe 691
Db 2279 GTCGTCTCGATGATCTCTCTGGAGTGTGGCACTCGAAGAACAGCAGAAAGTTAAATC 2338
QY 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711
Db 2339 AGAGTGTGTCTTAAAGAACCTTGGACCATGACAGATGTGACACGACACACAGATTGTTA 2398
QY 711 rSerCysThrAlaAsnThr 717
Db 2399 CAGCTGCACAGCCAATACC 2417
```

```
US-09-893-238-10
; Sequence 10, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; CURRENT FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-10
Alignment Scores:
Pred. No.: 2,18e-110 Length: 1051
Score: 1226.50 Matches: 236
Percent Similarity: 84.45% Conservatives: 3
Best Local Similarity: 83.39% Mismatches: 31
Query Match: 15.50% Indels: 13
DB: 9 Gaps: 3
```

US-09-787-097-12 (1-1429) x US-09-893-238-10 (1-1051)

```
QY 2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAlaAla 21
Db 217 GTGGCGCGCGCGCGCGCGCTGAGCGCGCGCTGAGGGGAGGAGCAGCAGCAGCAGCGCG 276
QY 22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41
Db 277 CTGCGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 336
QY 42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56
Db 337 GGACCGCGCGCGCGCGCTGTCTCCCGCGCGCTGTCTCCCGCGCGCTGTCTCCCGCGCG 396
QY 57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 76
Db 397 CTGCTGCGG---CTGCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432
QY 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 96
Db 433 CTGCCCCGCGGAGCGCGCGCGCTGCGGTGGCGCGCGCGCTGCGGTGGCGCGCGCGCGCG 492
QY 97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
Db 493 GAGGCCAAGAAATGTACCGCGCGCTGTGTCAACGCGCGCGCGCTGTCAACCTCGACCGCG 552
QY 117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
Db 553 CAGTGGCTGTCGCCACCGCGCTGGGTGGCGGAGCAATGCCAGCACTGCGGGGCGCGCTTC 612
QY 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156
Db 613 AGACTAATCGCTCTCTGCAATTGTAAACAGATGGACCTGGGAATTAATAATATAAGACG 672
QY 157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db 673 AAGTGCACATGGCTCTATTGAAGACAGACCAATAGAAATATGAGACTTCGCTTCAACCAT 732
QY 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
```



Db 733 TTGTCTACAGATGATGGACCACTTATATATGTTTATGATGGGACTCAATCTACGCA 792  
Qy 197 ProLeuValAlaAlaPheSerGlyLeuileValProGluArgAspGlyAsnGluThrVal 216  
Db 793 CCTCTGATGCTGCTTTAGTGGCTCATTTGTTCTGAAAGAGATGGCAATGAGACGGCT 852  
Qy 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyr 236  
Db 853 CTTGAGGTCACCTGCTCACTTCACTGATGATGATGCTGCTGCTTTCAGTATGCTGCTTAT 912  
Qy 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256  
Db 913 AATCTGACTGGATTTAATATCACTTACAAATTTTGACATGCTGCTGCTGCTGCTGCTGCT 972  
Qy 257 ArgGlyGluCysLeuLeuSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276  
Db 973 CAGAGAGAGTGAAGAGCAGTACAGCAGCAGCGCTGTGTGATGTGAATGCTTGAAGAAC 1032  
Qy 277 TrpLysGly 279  
Db 1033 TGGAAAGG 1041

## RESULT 9

US-10-085-198-15  
; Sequence 15, Application US/10085198  
; Publication No. US20040009907A1  
; GENERAL INFORMATION:  
; APPLICANT: Alsobrook et al.  
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same  
; FILE REFERENCE: 21402-279  
; CURRENT APPLICATION NUMBER: US/10/085,198  
; PENDING FILING DATE: 2002-02-25  
; PRIOR APPLICATION NUMBER: 60/271,646  
; PRIOR FILING DATE: 2001-02-26  
; PRIOR APPLICATION NUMBER: 60/276,401  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/311,981  
; PRIOR FILING DATE: 2001-08-13  
; PRIOR APPLICATION NUMBER: 60/312,858  
; PRIOR FILING DATE: 2001-08-16  
; PRIOR APPLICATION NUMBER: 60/271,840  
; PRIOR FILING DATE: 2001-02-27  
; PRIOR APPLICATION NUMBER: 60/277,324  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/286,096  
; PRIOR FILING DATE: 2001-04-21  
; PRIOR APPLICATION NUMBER: 60/299,695  
; PRIOR FILING DATE: 2001-06-20  
; PRIOR APPLICATION NUMBER: 60/315,614  
; PRIOR FILING DATE: 2001-08-29  
; PRIOR APPLICATION NUMBER: 60/272,405  
; PRIOR FILING DATE: 2001-02-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 653  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 15  
; LENGTH: 8589  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-085-198-15

Alignment Scores:  
Pred. No.: 9,966-83 Length: 8589  
Score: 960.50 Matches: 333  
Percent Similarity: 36.49% Conservative: 152  
Best Local Similarity: 25.06% Mismatches: 496  
Query Match: 12.14% Indels: 350  
DB: 15 Gaps: 55

US-09-787-097-12 (1-1429) x US-10-085-198-15 (1-8589)

Qy 104 ProCysValAsnGlyGlyArgCysAsnProGly-----ThrGlyGlnCysValCysPro 121  
|||||

Db 61 CCGTGC-----TGGGGTCCGTCTCC-CCTGGGGCCCGGGGGGAGCTGCAAG----- 107  
Qy 122 AlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySer 141  
Db 108 -----GGGCAGCGGAGGTGCTGCGGAGGCG 134  
Qy 142 SerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeu 161  
Db 135 CAGGCTTCGTGACGATGTCGGGCACTCAGCGTCAATGGCACTCGGAGTGCTC 194  
Qy 162 IleGluGly---GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180  
Db 195 ATCGAGGCCCCAAGCCCGCAGCACCGGATCCTCTGGACTTCCTTTCTTCTGGACACAGAG 254  
Qy 181 CysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200  
Db 255 TGACGATATACACTCTGTTGTGATGACGGTACTCCCGCGAGGGCGCTGCTGTGC 314  
Qy 201 AlaPheSerGlyLeuileValProGluArgAspGlyAsnGluThrValProGluValVal 220  
Db 315 AGTCTAAGTGGGAGCACCCGACCT-----CGCCCATCGAA 350  
Qy 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240  
Db 351 GCTTCTCAGGCAAGATGCTGCTGCACCTCTTCAGTGATGCCAACCTACAACTGCTGGGC 410  
Qy 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260  
Db 411 TTTAAGCCTCATTCGCTTCTCCCTGTCGGGGTGGTGCAGAGCCACGGGCAAGTGC 470  
Qy 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280  
Db 471 CAGCCACCGGGT-----GTGTGTGCTTGGAGCGGGCTGGGGGGTCTCT 515  
Qy 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300  
Db 516 GACTGTGGCCTGCAGGAGTCTCAGCCTACTGTGGC-----AGCCACGCGACCTGC--- 566  
Qy 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320  
Db 567 GCCTCGCCCTGGGACCATCGCTGTGAGCCTGGCTTCTTGGGAGCTGCTGTGACCTG 626  
Qy 321 ProValProAlaAsnGln-----SerPheTrp-----ThrArgGluGluTyr 334  
Db 627 CACCTGTGGGAGAACACAGGGGGCTGGGTGTGTGGCACACAGTGTAGTCCAGGAGCCCTGCC 686  
Qy 335 SerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrp 354  
Db 687 TTCTCTGCCGTATTGGGGCAGCTGGCGCTTCTCTCCCCACCAGGG---CTGCTGGCA 743  
Qy 355 ValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeu 374  
Db 744 GTTTTCGAGGCCAGGACCTCAACATGCTGGGTGACCTGCTCTA---TACAACTTC 800  
Qy 375 AlaserArgGluTrpLeuProLeuAsnArgSerValAsnValValValArgTyrGly 394  
Db 801 TCCGCCAACACCTGGGAGTCTTGGGACCTGAGT---CCTGCCCGGCTGCCGTCTACTCC 857  
Qy 395 HisSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyLysIleAspProThr 414  
Db 858 CATGTGGCCGTGGCCTGGCGCGCTCCCTGGTACTGATGGTGGTGAGCTG---GCTGAC 914  
Qy 415 GlyAsnValThrAsnGluLeuArgValPheHis---IleHisAsnGluSerTrpValLeu 433  
Db 915 GGCTCGCTCACCAACAGCAGTGTGGGCTTTCAGTCCACCTGGGAGGGGCCACCTGGGAGCTC 974  
Qy 434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThr 453  
Db 975 CTGGCACCACTGCTCCAGCTCCTCGGGGGCCCCAGGCTCGCAGGTCAACGGGCTGCC 1034  
Qy 454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473  
Db 1035 CTGGTGGATGAT---GTCTGGCTATATGTGTCTGGAGGCGCGACCCCGCCAGCACTCTTC 1091

QY 474 IleSerAsnValGlnGluTyrAspLeuAspLysAsnThr-----TrpSerIleLeu 490  
DB 1092 TCCTCTGGCCTCTTCGCTTTCGGCTTGACACACAGCGGGGGCTATTGGAGCAGGTG 1151  
QY 491 HisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSerSerValTyrAspHisArg 510  
DB 1152 ATTCGGGAGGCGGAGCGGCCCTCTGCTCCACTGCGCCACTCCATGGTGTTCATGCCGCC 1211  
QY 511 ThrArgAlaLeuTyrValHisGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeu 530  
DB 1212 TCCCGTGCCTGCTGCTCATGTGTGACACACCGGCCCTCCACTGCGCGTTCCTGTGCGA 1271  
QY 531 AlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTyrThrIleLeuLysAspSer 550  
DB 1272 GTGAACCTCCACTGAGCTTTTCCACGTGGATCGGCATGTGTGGACGACGCTGAAGGGCGG 1331  
QY 551 ArgPhePhe-----ArgTyrLeuHisThrAlaValIleValSerGlyThr 565  
DB 1332 GATGGCTTCAGGCGCCCAAGGAGCGAGCCCTCCACACAGCCAGTGTTCGGCAATTAC 1391  
QY 566 MetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCys 585  
DB 1392 ATGGTGTCTATGGGGGCAATGTGCACACCCCAATACAGGAGGAA-----AAGTGC 1442  
QY 586 PheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTyr----- 600  
DB 1443 TACGAAGATGGCATCTTCTTACCATCTTGCTGCCATCATGGTGTACAGGAGCTCAG 1502  
QY 601 -----SerValLeuProArgProAspLeuHisAspValAsnArgPheGlyHis 617  
DB 1503 CTTCGCCCGCCAGGAACCCCTGAGGCGGAGCAGCGCCCTCCAGTGTTCGTACTCATCAT 1562  
QY 618 ---SerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
DB 1563 GTAGCTCGCGTCTTGGTGGCAGCTCTGTGTGGTGGTGGGTGTACAGCGCGCGGCC 1622  
QY 637 LeuSerAspIleLeu-----ValPheThrSerGluGlnCysAsp 649  
DB 1623 CTGGGGACTTACGGCTACAGGTGCCCTTTGTGTTCAGGCACCTGCCCTGAC 1682  
QY 650 AlaHisArg-SerGluAlaAlaCys-----LeuAlaAlaGlyProGlyIleArg-- 665  
DB 1683 TACCACCTTGACTACTGTCTCCATGTATACACACACAGCGTCTGCTCCCGGACCCGGAA 1742  
QY 666 CysValTyrAsnThrGlySerGlnCysIleSerTyrAlaLeuAlaThrAspGluGln 685  
DB 1743 TGCAGTTGTGCCAGGAGCCTGCCAA-----GCTGCACCCCTCTCT 1784  
QY 686 GluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAsp 705  
DB 1785 GGGACCCCTTG---GGGGCTTGTCCAGCGCCAGCTGCTGGGCTGGGCGGCTC--- 1838  
QY 706 GlnHisThrAspCysTyrSerCysThrAla-----Asn 716  
DB 1839 ---CTGGGTGACTGCCAGCGCTGCTGCTTTCAGAGCCCAAGCCCTCCAGCGGGA 1895  
QY 717 ThrAsnAspCysHisTrpCys-----AsnAspHisCysValProArgAsnHisSer--- 733  
DB 1896 CTTGGCACCTTGGGCTGTGTGTGTCGACATGAGAGTGTCTCTTCTAGGCTGAGCAGGCC 1955  
QY 734 ---CysSerGluGlyGlnIleSer-----IlePhe 742  
DB 1956 CGTCTCGGAGGAGGAGCAGATCTCAGGCACTGTGGGCTGTGGGGCGCTGCGCTGTCTTC 2015  
QY 743 ArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys-AsnLysLysThrSerCys 762  
DB 2016 -----GTCAGTCTCCCTGGAGCGCTGGTCAACCCAGAGCTTCTCTG 2054  
QY 762 sArgSerCys----- 765  
DB 2055 CTGGCGCTGACCTTGCTACCTTTTCAGCAGCGCCCAATACCTCCAGCGCTGACAAGTCT 2114

QY 766 -----AlaLeuAspGlnAsnCysGlnTr 773  
DB 2115 TCAATTGTCCGAGCAGACACCATCACCTTAACACCAGCGCAGACAGATGTGTCCCTG 2174  
QY 773 pGluPro-----Ar 776  
DB 2175 GTCTACCGTGGCTTCATCTACCAATGTCTCCTGGAGGCGCAGGTGGACACAGGGCTGAG 2234  
QY 776 gAsnGlnGluCys-----IleAl 782  
DB 2235 GAGCTGCGCGTGTGGACGCGGCGCCAGCCCTACAGTCTCTGGCCCGAGATGCCGTGCG 2294  
QY 782 aLeuProGluAsnIle-----CysGlyIleGlyTrp----- 792  
DB 2295 CTTGACACGGAACATGAGGAGGTTGGGCGCTGGTGGCTCATCAGGAGAAGGAGACG 2354  
QY 793 -----HisLeuValGlyAsnSerCysLeuLysIleThrAl 805  
DB 2355 CGCGCGCTGCAGCGCCCTGGGTGTGTCTCCCTTCTCCCTGCTGGCGGAGCACAG 2414  
QY 805 aLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuAla-- 824  
DB 2415 TATGAGTAGAGATCAGGCGCCAGCTCAATGCTCGCAGGCGCTGGGACAGCGAGCTA 2474  
QY 824 ----- 824  
DB 2475 ACTCTGCTGGGATCGGACTGGTGTGCCAGGAGCGAGATCTCTTCTTCTCTG 2534  
QY 825 -SerLeuThrThrGlnLysLysValGluPheValLeuLys-----GlnLe 839  
DB 2535 GAGCCCTACCGTCGCTGCTGACCTCTATCTTCTGCTGGGTGCTGTGGCAGAC 2594  
QY 839 uArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThr-----ProTr 856  
DB 2595 CAGGGCTGTGGCTGGGCTGACCTGCACCTGCCTGCCAGCGCGGCGGAGCCAT 2654  
QY 856 pValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheh 876  
DB 2655 TGGCGGGATGACGGGCTGGTGGTCTCTGCTGGTCTGG-----TGCTTACCTTCTG 2708  
QY 876 rAsnSerLeuLeuGlnTrp-----MetProSerGluProSerAspAgl 891  
DB 2709 CCACTCTGCAGGAGCATCGGACTGCCAGCTGCACCCAGAGCCCTTCTGTAGTGG 2768  
QY 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911  
DB 2769 C-----ATCAGAGCACCGCCGCAAAAGGGGACGCGGCATGCAGCC----- 2808  
QY 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgth 931  
DB 2809 -----GGCGGGGCGG--GGTGGGGTGGCTTGAAGATCCAGAGAGTGTCCCC 2857  
QY 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951  
DB 2858 GCTCTGCAGCAGCAGCTGACCTGTGAGGACTGCTGGCCAACTCTAGCAGTGGCCCTG 2917  
QY 951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPheProPheG 971  
DB 2918 GTGCCAGTCCACCCACACCTGCTCTCTGTTGCTGCTACTTGGCCCGGTACCCACAG 2977  
QY 971 yGlnCysMetGluTrp-----TyrThrMetSerThrCysProProGluAsnCy 987  
DB 2978 GGGCTGTGAGGCTGGGACGACAGTGTACATCGAGCCACCGTGC-----CGGAGCTG 3031  
QY 987 sSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAs 1007  
DB 3032 CGATGGCTTCTCCTCCTGATGAGTGTCTGACAGCCAGGAGTGTGGTGTGTGGCAA 3091  
QY 1007 pProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro----- 1024  
DB 3092 TGAGGACACCCACACATGGGAGCGGTGCTTACAGGGGGGACTTCTCAGGGCCCTCGGTG 3151  
QY 1025 -----ValLysMetProSerGlnAlaProThrGl 1034

```
Db 3152 GGGTAACAGTCTCCCTGGTGGGGAGGCGCTGGGGCTTCCCTGGCCCTCCCTGCCCG 3211
Qy 1034 yAsnPheTyrProGln-----:::---:::---|||
Db 3212 CTGGGCATACGCCCGCTCTCCAGCTGGATGATGCTGCCCTGGGCCCTGGCCGGTGCCA 3271
Qy 1040 -----ProLeuLeuAsnSerMet-----
Db 3272 CCGCGGGGCGACCTCCCTGAACACGCCCTCAGTACGAGTGTCACCTGCCAGCGGGGCTA 3331
Qy 1046 -----
Db 3332 CAGGGTATGTCATCTCACCTGCACCGCACCGATGGCATCTCACACTGCACACCGCAC 3391
Qy 1047 -CysLeuGluAspSerArgTyr-----
Db 3392 GTGCTTGAGGACTGTGGCATGTGTGTGACGTGGCGCCCGGACTTTACCTGGGTGTG 3451
Qy 1054 -----AsnTrpSerPheIleHisCysProIle-----
Db 3452 TGACCTAGGCTGGACATCAGACTGCCCTCCCTCCACACCGCGCGGTGCCCGACGCC 3511
Qy 1063 -----CysGlnCysAsnGlyHisSerLysCysIleAsnGln-----Se 1075
Db 3512 CGCTGCTCCCGGACTGTGGCTGCAGCTTCCACAGCCACTGCCCAAGCGGGCCCTGG 3571
Qy 1075 rIleCysGluLysCysGluAsnLeuThrGlyLysHisCysGluThrCysIleSerG 1095
Db 3572 CTTCTGCACGAGTCCAGGACTGGACATGGGGGAGCACTGCCAAGCATGCCGCGCGG 3631
Qy 1095 yPheTyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHis 1115
Db 3632 CAGCTTCGACAGCCACAGGCTTAGGGGCTGGCGCCCTGCCAGTGCAACGGCACGG 3691
Qy 1115 a-----SerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysG 1131
Db 3692 GGACCCACGCCCTGGCCACTGGGACACCTCAGTGGGCTGCTCTCTGCCAGGACCA 3749
Qy 1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnPro 1150
Db 3750 -ACGAGGGTGGCCACTGCCACTGCG-----TCCCGAGCTATTATGGGGATCCCG 3802
Qy 1151 -LeuArgGlyThrCysTyr 1156
Db 3803 GGCGGTGGTTCCTGCTTT 3821
```

## RESULT 10

```
US-10-120-988-143
; Sequence 143, Application US/10120988
; Publication No. US20030219745A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Goodrich, Kyle
; APPLICANT: Liu, Chenghua
; APPLICANT: Ren, Feiyun
; APPLICANT: Wang, Dunrui
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219745A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 802CON
; CURRENT APPLICATION NUMBER: US/10/120,988
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: 09/774,528
; PRIOR FILING DATE: 2001-01-30
; NUMBER OF SEQ ID NOS: 441
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 143
; LENGTH: 7244
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
```

```
; LOCATION: (714)..(6518)
US-10-120-988-143
Alignment Scores:
Pred. No.: 3,59e-79 Length: 7244
Score: 923.00 Matches: 423
Percent Similarity: 32.95% Conservatives: 206
Best Local Similarity: 22.16% Mismatches: 576
Query Match: 11.66% Indels: 706
DB: 15 Gaps: 92

US-09-787-097-12 (1-1429) x US-10-120-988-143 (1-7244)
Qy 2 ValAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAlaAla 21
Db 1279 GTCACTCCACGCGGGGTACC-----AGGGTATGTCATCTCACACTGCACCGCACGT 1332
Qy 21 aLeuAlaGlyArgSerGlyGlyProHisTrpAspValThrArgAlaGlyArgPr 41
Db 1333 GCTTGGAGGACTG-----TGGCCATGGTGTGTG-----CAGTGGCCCC 1371
Qy 41 oGlyLeuGlyAlaGlyLeuArgLeu---ProArgLeu----- 52
Db 1372 CGGACTT-----TACCTGGTGTGTGACCTAGGCTGGACATCAGACCTGCCCTCCCA 1425
Qy 53 -----LeuSerProProLeuArgProArgLeuLeuLeuLeuLeuLeuLeu 67
Db 1426 CACCGCGCGCGGTCCGCCAGCCCCCGCTGCTCCCGGA-----CTGTGGCTGCA 1476
Qy 67 uLeuProProProLeu-----LeuLeuLeuLeuLeuProCysGluAl 81
Db 1477 GCTTCCACAGCCACTGCCGCAAGCGGGCCCTGCTCTGCGACGAGTGCAGG-ACTGG 1535
Qy 81 aGluAlaAlaAlaAla-----AlaAlaValSerGlySerAlaAlaAlaGluAla 99
Db 1536 ACATGGGGGAGCACTGCCAAGCATGCCGCCCGCAGCTTCGGCAACGCCAGGCTCT 1595
Qy 99 sGluCysAspArgProCys---ValAsnGly-----GlyArgCysAs 112
Db 1596 AGGGGCTGC-CGGCCCTGCCAGTGCACCGGACAGGGGACCCACGCGGTGCCACTGCGA 1654
Qy 112 nProGlyThrGlyGlnCysValCysProAlaGlyTrpValGlyGluGlnCysGln 130
Db 1655 CAACCTCAGTGGGTCTGCTTCTGCCAGGACACACCGAGGGTGCCTCCACTGCCAGTCTG 1714
Qy 131 -----HisCysGlyG 134
Db 1715 CTCCCCAGGCTATTATGGGGATCCCGGGCGGTGTTCTCTGCTTCCGGGAGTGTGGGG 1774
Qy 134 yArgPheArgLeuThrGlySerSer----- 142
Db 1775 TCGCGCCCTCTCTCACCAACGTGCTCTCAGTGGCAGTGGGCTCAGCGGGTCCGGGGCT 1834
Qy 143 -----GlyPheValThrAspGlyProGlyAsnTyrLysThrLys 157
Db 1835 GCTGCTCCAGGTGGGGGCTGCAAGAGCGGGCTGCG-----CTGCTCTA 1882
Qy 157 sCysThrTrpLeuIleGluGly-----GlnProAsnArgIle-- 169
Db 1883 CTGTGTGGGTGTCTCGGCCACTGAGAGCTACAGCCCTGTGCTCCCGGACCTCTG 1942
Qy 170 -----MetArgLeuArgPheAsn---HisPheAlaThrGluCysSerTrpAspHis 186
Db 1943 TCCCCCACTCACCTCTACCTTCTCCCGCAGCAGCAGCAGCCCCCTGCACGCTGAGTACGT 2002
Qy 186 uTyrValTyrAspGlyAspSerIleTyr-----AlaPr 197
Db 2003 CTGGCGGTTCATGGATTCACCGCTCTCTGGACACTGGTGTGTTCAGTCCGACCGCAG 2062
Qy 197 oLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValPr 217
Db 2063 CCTCATAGCTGCTTCTGCGGC-----CAGCGACGCGGACAGCGCCCTCCTGTTCA 2113
```





```
QY 1298 aValValThrLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuLeuAr 1318
Db 5981 GCTCTCTCTGGAAGCCAGCAGGCTCTGGACCGGCGGAGGAGCAGCGCGCGCACTTCCA 6040
QY 1318 gLuuMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAs 1338
Db 6041 GGAGATGACCAAGATGCCAGCGCCCTTCGCCAAGGTACCGCTCTCTCCCAACCTGA 6100
QY 1338 p----- 1338
Db 6101 CCTACTGCTCCCGCTCCGCTGGAAGCGGCTGGGCTCCACCTCCCGCTTCGCGCC 6160
QY 1339 -GluGluPro-----ProAspLeuIleGlySer----- 1348
Db 6161 CTCTGAGCCCTCTCTGGCAGCCCTGCTGTGACAGGGGCGGCTGGGCGCCCGGACCCAT 6220
QY 1349 -----IleLysThrValProLys-----pr 1355
Db 6221 GSGAGGGGCTGCTGCCACCGACCCATCCCGCCACCACTGCTGGGCTGGGCTGGGCC 6280
QY 1355 oLleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLe 1375
Db 6281 CATCACTCTCGAGCCACAGAGATGGCATGGCTGGCGTGGCCACACACTGCTCTCCAGCT 6340
QY 1375 uProArgGlyLeuGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAl 1395
Db 6341 GCTGGCGGG-----CCCATGCAACCCAGCGCGCTGCTGGGCTCAGC 6385

RESULT 11
US-10-277-216-5
; Sequence 5, Application US/10277216
; Publication No. US20040002470A1
; GENERAL INFORMATION:
; APPLICANT: KEITH, TIM
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,
; TITLE OF INVENTION: OBESITY, AND INFLAMMATORY BOWEL DISEASE
; FILE REFERENCE: 2976-4051
; CURRENT APPLICATION NUMBER: US/10/277,216
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: 10/126,022
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 09/834,597
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 09/548,797
; PRIOR FILING DATE: 2000-04-13
; NUMBER OF SEQ ID NOS: 420
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 207433
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-277-216-5

Alignment Scores:
Pred. No.: 1,3e-68 Length: 207433
Score: 840.00 Matches: 231
Percent Similarity: 24.34% Conservative: 10
Best Local Similarity: 23.33% Mismatches: 34
Query Match: 10.61% Indels: 717
DB: 15 Gaps: 6

US-09-787-097-12 (1-1429) x US-10-277-216-5 (1-207433)

QY 796 GlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPhe 815
Db 7134 GTGGCTCATGCTATAT-ATCTCAACATTTGGGAGGCCAAGGTGGGAGGATCACTTGGAG 7192
QY 816 CysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheVal 835
```

```
Db 7193 CTGAG-----AGTTTGAACCTAGCCAGGGCAACATGGCAAGACTTCATCTCTCCAAAATT 7246
QY 836 ---LeuLysGlnLeuArgIleMetGlnSerSerGlnSerMet-----SerLys 850
Db 7247 TAAATTAATAAAGATACATCTGTATTATCATAGTTGTGCTCTTTTGGCTTTAGTCCAAG 7306
QY 851 LeuThrLeuThrProThrValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGlu 870
Db 7307 CTCACCTTAACCCCATGGGTGGCTTCGGAAGATCAATGTGTCTCTACTGGTGGCTGGGAA 7366
QY 871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
Db 7367 GATATGTCCCATTTTACAATAGTTTACTACAGTGGATGCCGCTCTGAGCCCATGATGCT 7426
QY 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
Db 7427 GGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTCAACCTGCATC 7486
QY 911 AsnProLeuAsnGlySerValCysGluArg----- 920
Db 7487 AACCCACTCAATGGTAGTGTCTGTGAAGGCCCTGGTAACTTACAGGTGAATTAGTGGT 7546
QY 920 ----- 920
Db 7547 ATTCAGAGTTTATTGTGAGAGAAACCATAGGAGCATAGTTCAATTCCTGAGATGTGTGAA 7606
QY 920 ----- 920
Db 7607 GTAGTCATGAAAACAGATGAAGTATTGATTCAAGCATGCAAGAGAGATATAACCCAGA 7666
QY 920 ----- 920
Db 7667 TTTCAGACGACGAGGAATATTCTGGACCCCTGAATAGTTTAAATATAAGCAAACTA 7726
QY 920 ----- 920
Db 7727 AAAATAACTAACTACTCGAAGAACTGATATTCTAATTAACAATGAGATTGATAGTT 7786
QY 920 ----- 920
Db 7787 TATTGACCAGAAAAGTATTGAGAAATTGCTCTGAAGCAAAATTTATTTGGTGTGSCAGA 7846
QY 920 ----- 920
Db 7847 GAAATGCTGTGGAAGAGAAACAAAGAAAGGAAATATAAACCAAGAAAGATTAGTAAAG 7906
QY 920 ----- 920
Db 7907 CAAGCAAGTGACTGTCAGGACAGTGTTCAGAAGGTAGTGTCAACAGGGAGAGAAAATGAT 7966
QY 920 ----- 920
Db 7967 GAGAGCAGTTCTGTAAGCGAGGACAGACCAACCCAGAACTAGCGGAGAGCTGCAGGAGA 8026
QY 920 ----- 920
Db 8027 GTACTGAGCAGACAGACACTCGAGTAGTCTTCCAGCTTTTCACTCTCTGGGCTACATT 8086
QY 920 ----- 920
Db 8087 TTGGCTAACTAAGGAGGCCCGGAGGTAGTCTGTCAGCATGAGACAACTCAAAATAA 8146
QY 920 ----- 920
Db 8147 GCCTCTCTCTCTGACAGGAGACTCAAGAGCAGTGGTGAATTTGGGTATTATTTGGGA 8206
QY 920 ----- 920
Db 8207 GGCTGGACTCCATTTGCTATTTGCTCCAAATTTGGGAAAGTAAATTTCTCTGAGGAATTACA 8266
QY 920 ----- 920
Db 8267 GAGGTATAGGAGAGAAATTCATAGCATTTGTAGATTTTCTAAATATTGATTTCTAAACTCTC 8326
```

QY 920 ----- 920  
Db 8327 TAGGGAGCTCCCGAGAGCCTTTTCAGAAATTCCTGAATCCCTTTGTTGAACCTTAATGAATGC 8386  
QY 920 ----- 920  
Db 8387 CGTAGACCTTCTCCCTCCAAATTTGCAACATGAGATTTTACATATTCAAGTGGATTGT 8446  
QY 920 ----- 920  
Db 8447 AAAACCCCTAAAGTTCATCTGGGTCCCAAAATTCCTAAGGGGCTTACAGCCCTACTTTTA 8506  
QY 920 ----- 920  
Db 8507 TGGAAAGATTCACTGTTTCTCTCATCTTTGCTCTTAATGCTCTTTGAAAGAAAATTTATT 8566  
QY 920 ----- 920  
Db 8567 TTTTCCACAAGTGTAAATACACTTGACTAAACAGTACTTGATGACTTTTATTGTTTTC 8626  
QY 920 ----- 920  
Db 8627 TATTTCTCTCATTTAATTGATTTTCATGGCACATTAATGGTGCATCACCACATTTGAAA 8686  
QY 920 ----- 920  
Db 8687 AGTCTTGCTGGCACAGTGGCTCAGCCTGTAAATCCAGCACTTTAGAGGCCAAGACG 8746  
QY 920 ----- 920  
Db 8747 GGTGATCACAGGTCAAGATATCAAGACCATCTGGTAAACATAGTGAACCCCATCTC 8806  
QY 920 ----- 920  
Db 8807 TACTAAAAATACAAAAATTAGCCAGAGCTGATGGCGCACCCCTGTAGTCCAGCTACTT 8866  
QY 920 ----- 920  
Db 8867 GGGAGGCTGAGGAGGAGATCACTTGAAACGGGGAGCGGAGGTTCAGTGAAGCCGAAA 8926  
QY 920 ----- 920  
Db 8927 TTGTGCCACTGCACCTCTAGCTGGAGACAGAGTCCGCTCTCAAAAAAAGAA 8986  
QY 920 ----- 920  
Db 8987 AAAGAAAGAAAGAAAGCTTTTGACCTTAGCGGACATGGTGGGAGCTCTAAGTGTCTCTC 9046  
QY 921 -----ProAlaAsnHisSerAlaAlaGlnCysArgThrProCysAlaLeuAr 936  
Db 9047 TTGGTTTCATTCAGAGAAACCAAGTCTAAGCAGTGGCGACACCAATGTGCTTGAG 9106  
QY 936 gThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLy 956  
Db 9107 GACAGCATGTGGAGATTGCACAGCGGAGCTCTGAGTGCATGTGGTGCAGCAACATGAA 9166  
QY 956 sGlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTr 976  
Db 9167 CGAGTGTGGACTCAATGCTAGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATG 9226  
QY 976 pTyrThrMetSerThrCysProPro----- 984  
Db 9227 GTATACGATGACACCTGCC-CCGTAAGTGAAGGAGGAGCCCTAGGCACTTTATGCATG 9285  
QY 984 ----- 984  
Db 9286 CCCCTGTATAGGCAACAACTCAGCCATGAGGCTGTGTGTCAGCCTCTGAACATTTAG 9345  
QY 984 ----- 984  
Db 9346 AAACAGACTGGACATGACCTCTGCTCAACCTGACACAGACTGCCATCGAGACCTTGC 9405

QY 984 ----- 984  
Db 9406 TGCCTATTGAGAACCTTTCATACAGAATCAGGCACATTCAGACTAAATAAATGTAAGATAG 9465  
QY 984 ----- 984  
Db 9466 ATCACAGAGTACAGAAATAACTTTGTCCAACTTCAGTGTGCATATTGCTCAATCCATGTAAT 9525  
QY 984 ----- 984  
Db 9526 ATCTCCATATCTGAATTCCTTAATTTGTAAAGTAAATGCTTTTCCAATAGATAATCTCTAA 9585  
QY 984 ----- 984  
Db 9586 GGTCCCTTTTGGCTTTCACATCTCGGATTCGAGAGAGGAGGAGGCTCATCTCTGTATT 9645  
QY 984 ----- 984  
Db 9646 GTATTGGGCAAAATACTGGGCTCTTTTACATTCATTCCTTTTAAATCAATCAAGACAGA 9705  
QY 984 ----- 984  
Db 9706 ATAAATATTTTGGACTCAAGCCAGTTTGATAGTCTGTAAATAAAGTAAATACAGTGAA 9765  
QY 985 -----GluAsnCysSe 988  
Db 9766 TTCAATCTACCTGTGATAGTCAATTCGAACCTTTTAAATAGCTGAAATTTGTTTC 9825  
QY 988 rGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyTyrCysThrAspPr 1008  
Db 9826 AGGCTACTGTACTGTAGTCAATTCCTGGAGCAACCCAGGCTGTGGTGGTACTGATCC 9885  
QY 1008 oSerAsnThrGlyCysGlyCysIleGluGlySerTyrLysGlyProValLysMetPr 1028  
Db 9886 CAGCAATATCTGGCAAGGAAATGATAGAGGTTCTTAAAGGACCCAGTGAAGATGCC 9945  
QY 1028 oSerGlnAlaProThrGlyAsnPhetyrProGlnProLeuLeuAsnSerSerMetCysLe 1048  
Db 9946 TTGGCAAGCCCTACAGGAAATTTCTATCCAGAGCCCTGCTCAATTCAGCATGTGTCT 10005  
QY 1048 uGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHi 1068  
Db 10006 AGAGGACAGACAGATACAACTGGTCTTTTCACTTCACTGTCCA-----GGTAA 10050  
QY 1068 sSerLysCysIleAsnGlnSerIleCys 1077  
Db 10051 GATGCTTGGCATATCCAAATTCAGTGT 10078

## RESULT 12

US-10-126-022-5  
; Sequence 5, Application US/10126022  
; Publication No. US20040023215A1  
; GENERAL INFORMATION:  
; APPLICANT: KEITH, TIM  
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,  
; TITLE OF INVENTION: OBESITY, AND INFLAMMATORY BOWEL DISEASE  
; FILE REFERENCE: 2976-4039US2  
; CURRENT APPLICATION NUMBER: US/10/126,022  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 09/834,597  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 09/548,797  
; PRIOR FILING DATE: 2000-04-13  
; NUMBER OF SEQ ID NOS: 420  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 207433  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-126-022-5

Alignment Scores: 1.3e-68 Length: 207433  
Pred. No.:



Score:	840.00	Matches:	231
Percent Similarity:	24.34%	Conservative:	10
Best Local Similarity:	23.33%	Mismatches:	34
Query Match:	10.61%	Indels:	717
DB:	16	Gaps:	6

US-09-787-097-12 (1-1429) x US-10-126-022-5 (1-207433)

Qy	796	GlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAenAlaLysLeuPhe	815
Db	7134	GGTGGCTCATGCCCTATATA-ATCTCAACATTTTGGGAGGCCCAAGGTGGGAGGATCACTTTGAG	7192
Qy	816	CysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheVal	835
Db	7193	CTCAGG-----AGTTTGAACACTAGCCAGGGCAACATGGCAAGACTCCATCTCCAAATTT	7246
Qy	836	---LeuLysGlnLeuArgIleMetGlnSerSerGlnSerMet-----SerLys	850
Db	7247	TAAATTAATAAAGATACTATCTGTAATTCATAGTTGTGTCTCTTTTGCCTTTAGTCCAAAG	7306
Qy	851	LeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGlu	870
Db	7307	CTCACCTTAACCCCATGGCTGGCGCTTCGGAAGATCAATGTGTCTACTGTGTCTGGTGGAA	7366
Qy	871	AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla	890
Db	7367	GATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGCTGTGAGCCACGATGCT	7426
Qy	891	GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle	910
Db	7427	GGATTCGTGGAAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGTGCGAACCTGCATC	7486
Qy	911	AsnProLeuAsnGlySerValCysGluArg-----	920
Db	7487	AACCCACTCAATGGTAGTGTCTGTGAAGGCCCTGGTAAAGTTTCACAGGTGAATTAGGTGGT	7546
Qy	920	-----	920
Db	7547	ATTCAGAGTTTATTTGTGAGAGAACCATAGGAGGCATAGTTTCATTTGCTGAGATGTGTGAA	7606
Qy	920	-----	920
Db	7607	GTAGTCATGAAACACAGATGAAGTATTGATTTCAAGCATGCAAGAAGAGTATAACCCAGA	7666
Qy	920	-----	920
Db	7667	TTTCAGAAAGCAGAAGGAATATTTCTGGGACCTGAAATAGTTTTTAATTATAAGCAAAACTA	7726
Qy	920	-----	920
Db	7727	AAATAACTAACACTACTCGAGNAAC TGATATTCATAATTAACTAGATTTGATAGTTT	7786
Qy	920	-----	920
Db	7787	TATTTGACCAGAAAAGTATTGAGAAATGCTCTGAAAAGCAAAATTTATTGGTGTGGCAGA	7846
Qy	920	-----	920
Db	7847	GAAATGCTGTGGAAGAAGAAACAAAAGGAAATAAATAAACAAGAAAAGATTAGTAAAG	7906
Qy	920	-----	920
Db	7907	CAAGCAAGTGACTCCAGGGACAGTGTTCAGAAAGGTAGTGTCAACAGGGAGAAAATGAT	7966
Qy	920	-----	920
Db	7967	GAGAGCAGTTTCTGTAAGCGAGGACAGAAGCAACCCAGAACTACGCGAGAGCTGCAGGAGA	8026
Qy	920	-----	920
Db	8027	GTACTGAGCAGACAGACACTCGGAGTAGCTTCCAGCCTTTCAGTCTCTCTGGGGCTACATT	8086
Qy	920	-----	920

8087	Db	TTGGCTAACTAAAGGAGCGCCCAAGGGTAGCTGCTGCAGCATGACAGCAACTCAAAATAA	8144
920	Qy	-----	920
8147	Db	GCCTCCTCCTCTGCAGAGGGAAGCTCACAAGCAGTCGGTGGAAATTGTGGTTATTTTGGGA	8206
920	Qy	-----	920
8207	Db	GGCTGAGCTCATTTCATTTGTCTCCAAATTTGGGAAAGTAATTCTCTGAGGAATTACA	8266
920	Qy	-----	920
8267	Db	GAGGTATAGGAGAGAATTCATACATTTGTAGATTTCTCTAAATATTGATTTCTAAACTCTC	8326
920	Qy	-----	920
8327	Db	TAGGGAGCTGCCAGCAGCCTTTCAGAAATCTGAATTCCTTTGTTGGAATTAAATGATGC	8386
920	Qy	-----	920
8387	Db	CGTAGACCTTCTCCCTCCAAAATTCAAAACATGAGATTTTACATATTCAAGTGGATTGT	8446
920	Qy	-----	920
8447	Db	AAAACCCCTAAAGTTTCATCCTGGGTCCCCAAATTCAGGGGCTTACAGCCCTACTTTTA	8506
920	Qy	-----	920
8507	Db	TGGAAGATTCACTGTTTCTCTCATCTTTGTCCTTAATGCTCTTTGAAAGAAAAATTAIT	8566
920	Qy	-----	920
8567	Db	TTTTCCACAAGTGTGTAATACACTTGAATAACAGTACTTGATGACTTTTATTTGTTTC	8626
920	Qy	-----	920
8627	Db	TATTTTCTCATTTAATTGATTTTCATGGCACATTAATGGGTCAATCACCACCATTTGAAA	8686
920	Qy	-----	920
8687	Db	AGTCTTGGTGGCACAGTGGCTCAGCGCTGTAATCCACAGCACTTTTAGGAGGCCAAGACG	8746
920	Qy	-----	920
8747	Db	GGTGGATCACAGGTCAAGATATCAAGACCATCTCGCTAAATAGTAGTGAACCCCATCTC	8806
920	Qy	-----	920
8807	Db	TACTAAAAATACAAAAAATTAGCCAGACGTGATGGGCACCCCTGTAGTCTCCAGCTACTT	8866
920	Qy	-----	920
8867	Db	GGAGGCTGAGGAGGAGAAATCACTTGAACCGGGAGGCGGAGTTGCAGTAGCCGAAA	8926
920	Qy	-----	920
8927	Db	TTGTGCCACTGCATCTAGCCTGGAGACAGATGAGACTCCGCTCTCAAAAAAAGAAAGA	8986
920	Qy	-----	920
8987	Db	AAAGAAAGAAAGAAAGCTTTTGACCTTAGCGACATGTTGGGAGCTCTAAGTGTCTCTC	9046
921	Qy	-----ProAlaasnHisSerAlaLysGlnCysArgThrProCysAlaLeuAr	936
9047	Db	TTGGGTTTCATTTCCAGCAAAACACAGTGTCTAAGCAGTGCAGCACCATATGTGCTTGAG	9106
936	Qy	gthralaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLly	956
9107	Db	GACAGCATGTGGAGATTGACACAGCGGCAGCTCTGAGTGCATGTGTGACGAACAATGAA	9166
956	Qy	sglCysValAspSerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTr	976

Db 9167 CGAGTGTGGACTCCAAATGCTAGTGGCTCCTCCCTTTTGGCCAGTGTATGGAATG 9226  
Qy 976 pTyThrMetSerThrCysProPro----- 984  
Db 9227 GTATACGATGACACCTGCCCC-CCGTAAGTGAAGAGGAGCCCTAGGCACCTTATGCATG 9285  
Qy 984 ----- 984  
Db 9286 CCCTCTGTATAGGCAACAACCTCAGCCATGAGGCTGTGTGCAGCCCTCTGAACATTTTAG 9345  
Qy 984 ----- 984  
Db 9346 AACAGAGACTGACATGACCTCTGCTCAACCTGCAGAGAGACTGCCATCGAGACCTTGC 9405  
Qy 984 ----- 984  
Db 9406 TGCCTATTGAGAACCTTCATACAGAACTCAGGCACATTGACAGTAATAAATGTAAGATAG 9465  
Qy 984 ----- 984  
Db 9466 ATCAGAGTACAGAAATAACTTGTCCAACTTCAGTGTCTATTTGCTCAATCCATGTAAT 9525  
Qy 984 ----- 984  
Db 9526 ATCTCCATCTGAATTTCTTAATTGTAAAGTAAATGCTTTTCCAATAGATAAATCTCTAA 9585  
Qy 984 ----- 984  
Db 9586 GGTCCCTTTTGCCTTCAACATCTCGGGATTGAGAGAGGAGGGTCACTCTCTGTAT 9645  
Qy 984 ----- 984  
Db 9646 GTATTGGCAAAATACTGGGCTCTTTACATTCATTATCTCTTTTAAATAATCAAGACAGA 9705  
Qy 984 ----- 984  
Db 9706 ATAATATTTTACTCAAGCCAGTTGAATAGTCTGTAAATAAAGTAATAACAGTGAA 9765  
Qy 985 ----- 988  
Db 9766 TTCAGATCTACTGTGATAGTCAATTGCACTTTTATTTTATAGCTGAATTTGTC 9825  
Qy 988 rGlyTyrCysThrCysSerHisCysLeuGluInProGlyCysGlyTyrCysThrAspPr 1008  
Db 9826 AGGCTACTGTACTCTAGTCAATGCTGGAGCAACAGGCTGTGGCTGTGTACTGATCC 9885  
Qy 1008 cSerAsnThrGlyLysGlyCysLeuGluGlySerTyrLysGlyProValLysMetPr 1028  
Db 9886 CAGCAATCTGGCAAGGGAATGATAGAGGTTCTTAAGGACCAAGTGAAGATGCC 9945  
Qy 1028 oSerGlnAlaProThrGlyAsnPheTy-ProGlnProLeuLeuAsnSerSerMetCysLe 1048  
Db 9946 TTCGCAAGCCCTACAGGAAATTTCTATCCAGCCCTGTCTCAATTCAGCATGTGTCT 10005  
Qy 1048 uGluAspSerArgTyrAsnTyrSerPheHisCysProAlaCysGlnCysAsnGlyHi 1068  
Db 10006 AGAGACAGCAGATACAACTGCTCTTCAATCAGTGTCCA-----GGTAA 10050  
Qy 1068 sSerLysCysIleAsnGlnSerIleCys 1077  
Db 10051 GATGCTTGCATATCCAAATTCAGTGT 10078

## RESULT 13

US-09-893-238-5  
; Sequence 5, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238

; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 90050  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-5  
  
Alignment Scores:  
Pred. No.: 6,32e-62 Length: 90050  
Score: 766.00 Matches: 205  
Percent Similarity: 21.41% Conservative: 2  
Best Local Similarity: 21.20% Mismatches: 6  
Query Match: 9.68% Indels: 755  
DB: Gaps: 2

US-09-787-097-12 (1-1429) x US-09-893-238-5 (1-90050)

Qy 849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
Db 28996 TCCAAGCTCACTCTGACTCCATGGTTCCTTCGGAAGATCAATGTGTCTTACTGGTGC 29055  
Qy 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
Db 29056 TGGGAGGATATGTCCTCCATTCACAAATAGTTGCTGCATGGATGCCATCTGAGCCAGT 29115  
Qy 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908  
Db 29116 GATGCTGGCTTCTGTGGATCTTGTACAGAGCCTAGTACTCGGGATTAAGGCTGCAACC 29175  
Qy 909 CysIleAsnProLeuAsnGlySerValCysGluArgProAla----- 922  
Db 29176 TGCAATCAACCTCTCAATGGCAGCGTCTGTGAAGGCTGG-TAAGGACATGGGTGCATA 29234  
Qy 922 ----- 922  
Db 29235 TAGTGTCCAGGAGGAGCCCAAGACAGCAAGAGGACACAGCTGAATGAGGCTGAGGTGA 29294  
Qy 922 ----- 922  
Db 29295 TGAAGTACTTTATGCGCAGGAGGAGGAGCACCACATTTAGGCATATGTATTTCAAACAGA 29354  
Qy 922 ----- 922  
Db 29355 ACCCGATTCCAGATAGTCTTTCTTGGCCTCTGACTCTTTAAGCCATCTGAAACCAAA 29414  
Qy 922 ----- 922  
Db 29415 AATAAATTCTGAAAGAACCCAGTTTATATTGAGCTGCACCTGTTTCGTTGGTCTCAAAG 29474  
Qy 922 ----- 922  
Db 29475 TGTGAGAAATGTTCTAGAGATATTTCCTTGGTGTGTGCGAGAGAAGTGTCTATGGAGGA 29534  
Qy 922 ----- 922  
Db 29535 AACACAACTGAAACCAAGAAACATTTAGAAAAGCAGCAAGTCAGGACACTATTTCAGA 29594  
Qy 922 ----- 922  
Db 29595 CACTGTGGGTGGGGGAGGGGAGGCGATGSCCAAGAGCCGACAGAGCCACACCCAGGC 29654  
Qy 922 ----- 922  
Db 29655 TGTGGCAATGTCCTGGCTGAGGTTAAGGTTAGACTCCATGAGGCCAGGCCCAAGACAGC 29714



Search completed: March 1, 2004, 07:35:28  
Job time : 1542 secs

**This Page Blank (uspto)**

Result No.	Score	Query		DB	ID	Description
		Match	Length			
1	7696.5	97.2	859	3	US-09-245-041-14	Sequence 14, Appl
2	7303	92.3	8837	3	US-09-245-041-1	Sequence 1, Appl
3	6870.5	86.8	4072	3	US-09-245-041-16	Sequence 16, Appl
4	4136.5	52.3	2625	3	US-09-245-041-18	Sequence 18, Appl
5	3491.5	44.1	6370	3	US-09-245-041-12	Sequence 12, Appl
6	3435	43.4	2419	3	US-09-245-041-8	Sequence 8, Appl
7	1226.5	15.5	1031	3	US-09-245-041-10	Sequence 10, Appl
8	766	9.7	90050	3	US-09-245-041-5	Sequence 5, Appl
9	581.5	7.3	415	4	US-09-833-381-1916	Sequence 1916, Ap
c 10	454	5.7	17056	3	US-09-245-041-3	Sequence 3, Appl
c 11	446	5.6	246	4	US-09-702-705-966	Sequence 966, App
c 12	446	5.6	246	4	US-09-736-457-966	Sequence 966, App





QY 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756  
DB 2206 GCCAGATCTCCATTTTAGGTATAGAGAAATGGCCCAAGGATAACCCCTATGTAATCTGT 2265  
QY 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrrGluProArg 776  
DB 2266 AACAAAGAACCCAGCTGCAGGAGTGTGCCCTGGACACAGAACTGCCAGTGGAGGCCCGG 2325  
QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrrPheIleLeuValGly 796  
DB 2326 AATCAGGAGTGATGGCTCCCGGAAATATCTGTGGCATTTGGCTGATTTGGTTGGA 2385  
QY 797 AsnSerCysLeuLysIleThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
DB 2386 AACTCATCTTTGAAATTTACTTACTGCCAAGGAGAAATATGACAAATGCTAAATGCTTGT 2445  
QY 817 ArgAsnHisAsnAlaLeuAsnLeuSerLeuThrGlnLysLysValGluPheValLeu 836  
DB 2446 AGGAACCAATGGCTTTTGGCTTCTCTTCAACCCAGAAAGAGTAGAATTTGTGCTT 2505  
QY 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrr 856  
DB 2506 AAGCAGCTGCAGATATGCAATCTCTCAGAGCATGTCCAAGCTCACCTTAAACCCCATGG 2565  
QY 857 ValGlyLeuArgLysIleAsnValSerTyrTrrCysTrrGluAspMetSerProPheThr 876  
DB 2566 GTGGCCCTTCGAAGATCAATGTGTCTTACTGTGTCTGGGAGATATGTGCCCATTTACA 2625  
QY 877 AsnSerLeuLeuGlnTrrMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
DB 2626 AATAGTTTACTACAGTGGATGCCGTCTGAGCCCATGTATGTGATCTGTGGAAATTTTA 2685  
QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
DB 2686 TCAGAACCCAGTACTTCGGGGAGTGAAGCTGCAACCTGCATCAACCCACTCAATGTGTAGT 2745  
QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
DB 2746 GTCTGTGAAGGCCCTGCAACACAGTGTCAAGCAGTCCCGACACCATGTGCCCTTGAGG 2805  
QY 937 ThrAlaCysGlyAspCysThrSerGlySerGluCysMetTrrCysSerAsnMetLys 956  
DB 2806 ACAGCATGTGAGATTGCACAGCGGAGCTCTGATGTCATGTGGTCAGCAACATGAG 2865  
QY 957 GlnCysValAspSerAsnAlaTrrValAlaSerPheProPheGlyGlnCysMetGluTrr 976  
DB 2866 CAGTGTGTGGACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGG 2925  
QY 977 TyrThrMetSerThrCysProGluAsnCysSerGlyTrrCysThrCysSerHisCys 996  
DB 2926 TATACGATGAGACCTGCCCTCCCTGAAATTTGTTACAGCTACTGTACCTGTAGTCATTCG 2985  
QY 997 LeuGluGlnProGlyCysGlyTrrCysThrAspProSerAsnThrGlyLysGlyCys 1016  
DB 2986 TTGGAGCAACAGCGTGTGGCTGTGTGTACTGATCCAGCAATACTGGCAAGGGAATGC 3045  
QY 1017 IleGluGlySerTrrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036  
DB 3046 ATAGAGGGTTCCTATAAAGGACCACTGAAGATGCCCTTCGCAAGCCCTTACAGGAAATTC 3105  
QY 1037 TyrProGlnProLeuAsnSerSerMetCysLeuGluAspSerArgTrrAsnTrrSer 1056  
DB 3106 TATCCACAGCCCTCTCAATTCAGCATGTGTCTAGAGGACAGCAGATCACTGTGTCT 3165  
QY 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
DB 3166 TTCAATTCATGTCCAGCTTGGCAATGAACGGCCACAGTAAATGCAATCAATCAGAGATC 3225  
QY 1077 CysGlnLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096  
DB 3226 TGTGAGAGTGTGAGAACTTACCAACAGCAGAGCACTCGAGACTGTCATATCTGGCTTC 3285  
QY 1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116

DB 3286 TACGCTGATCCCAACCAATGGAGGAATGTGAGCCATGCAAGTGCATGGGCAAGGCTCT 3345  
QY 1117 LeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136  
DB 3346 CTCTGCACACCAACCAACGCGCAAGTCTTCTGACCACCAAGGCGCTCAAGGGGACGAG 3405  
QY 1137 CysGlnLeuCysGluValGluAsnArgTrrGlnGlyAsnProLeuArgGlyThrCysTrr 1156  
DB 3406 TGCAGCTATGTAGGTAGAAATCATACCAAGGAAACCCCTCTCAGAGAACATGTTAT 3465  
QY 1157 TyrThrLeuLeuIleAspTrrGlnPheThrPheSerLeuSerGlnGluAspAspArgTrr 1176  
DB 3466 TATACCTCTTATTACTATCACTATCAGTTTCACTTATGCTATCCAGGAAGATGATCGCTAT 3525  
QY 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196  
DB 3526 TACACAGCTATCAATTTTGTGGCTACTCTGTCAGAAACAAACAGGATTTGACATGTTTC 3585  
QY 1197 IleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrrAlaAlaSerPheSerAlaGly 1216  
DB 3586 ATCAATGCCCTCCAAGAAATTTCAACCTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGA 3645  
QY 1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTrrLysAsp 1236  
DB 3646 ACCCAGGCTGGAGAGAGATGCTGTGTTTTCAAAACCAACCAATTAAGAGATACAAAGAT 3705  
QY 1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTrr 1256  
DB 3706 AGTTTCTCTAATGAGAGATTTGATTTTCGCAACCCACCACCAATATCACTTCTTGTATT 3765  
QY 1257 ValSerAsnPheThrTrrProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPhe 1276  
DB 3766 GTCAGTAATTTTCACTTGGCCCATCAAAATTCAGATTGCCCTTCTCTCAGCACAGCAATTT 3825  
QY 1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuVal 1296  
DB 3826 ATGGACCTGTGTACAGTCTTCTGGACTTCTTTCAGTTGTTTCTCTCTTCTTCTCTGGTG 3885  
QY 1297 AlaAlaValTrrLysIleLysGlnSerCysTrrAlaSerArgArgGluGlnLeu 1316  
DB 3886 GCTGTGTGTGTTTGGAGATCAAAAGATTTGTTGGCTTCCAGACGTAGAGAGCAACTT 3945  
QY 1317 LeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGlu 1336  
DB 3946 CTTTCAGAGATGCAACAGATGGCCAGCTGCCCTTTGCCCTCTGTAAATGTGCTTGGAA 4005  
QY 1337 ThrAspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIle 1356  
DB 4006 ACAGATGAGAGCTCTCTGATCTTATTTGGGGGAGTATAAGACTGTTCACCAACCAT 4065  
QY 1357 AlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuPro 1376  
DB 4066 GCACCTGGAGCGGTGTTTGGCAACAAAGCGCTGTCTCTCTGTGTTTGTAGGCTCCCT 4125  
QY 1377 ArgGlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeu 1396  
DB 4126 CGAGGCTGTGGTGGCATCCCTCTCTGGGAGTCAAGTCTTCTGTGTGCCAGCGCTG 4185  
QY 1397 ValAspIleSerGlnGlnMetProIleValTrrLysGluLysSerGlyAlaValArgAsn 1416  
DB 4186 GTGACATTTCTCAGCAGATGCCGATGTGTACAGGAGAAAGTCAGAGCGGTGAGAAAC 4245  
QY 1417 ArgGlyGlnGlnProAlaGlnProGlyThrCysIle 1429  
DB 4246 CGGAAGCAGCAGCCCTCTCAGAGCTGGGACCTGCATC 4284

## RESULT 2

US-09-245-041-1

; Sequence 1, Application US/09245041

; Patent No. 6274339

; GENERAL INFORMATION:

; APPLICANT: Moore, K.



Db 1799 CAAGGCTCTGTACGTTTCATGGTGGCTACAAAGGCTTTACGGCCCAACAAATACCGGCTTGC 1858  
Qy 531 aAspLeuTyrArgTyrAspValAspThrGlnMetTrrThrIleLeuLysAspSerAr 551  
Db 1859 AGATGACCTCTACAGATACGATGGATGATCTAGATGTCAGACCACTTCTTAAGGACAGCG 1918  
Qy 551 gPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGI 571  
Db 1919 ATTTTTCCTTCTGTCATACAGCTGTGATAGTGTGAGTGAACCACTGCTGTGTTGGAGG 1978  
Qy 571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591  
Db 1979 GAACACACATGACACTTCCATGAGCCACGGTGCCTCAATGCTTCTCTCGGACTTCAT 2038  
Qy 591 tAlaTyrAspIleAlaCysAspArgTrrSerValLeuProArgProAspLeuHisHisAs 611  
Db 2039 GGCTTATGACATTTGCTGTGACCGATGCTCAGTGTCTTCCAGACCTGAGCTCCATCA 2098  
Qy 611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTrrValPheGlyGI 631  
Db 2099 TGTCACACATTTGGCCATTCAGCAGTCTTTGTACACACAGCACCATGTATGTGTTCGGCGG 2158  
Qy 631 yPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHi 651  
Db 2159 CTTCAACAGCCTCTCTCAGTGACGCTCTTGTCTTTTACCTCGGAGCAGTGCATGCACA 2218  
Qy 651 sArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgCysValTrrPheAsnThrGI 671  
Db 2219 CCGCAGTGAAGCTGTGTGTGACGACAGGACCTGGTATCCGGTGTCTGTGGACACACA 2278  
Qy 671 ySerSerGlnCysIleSerTrrPAlaLeuAlaThrAspGluGlnGluLysLeuLysSe 691  
Db 2279 GTCTGTCTGATGTACTCTCTGGAGTTGGCACTGAGACACAGCAGAAAGTTAAATC 2338  
Qy 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711  
Db 2339 AGAGTGTCTTTCTAAAGAACCTTTGACCATGACAGATGTGACACACACACAGATTGTTA 2398  
Qy 711 rSerCysThrAlaAsnThrAsnAspCysHisTrrPAsnAspHisCysValProArgAs 731  
Db 2399 CAGCTGCACAGCAATACCAATGATGTCACCTGGTGCAATGATCATCTGTGCTCCCTGTGAA 2458  
Qy 731 nHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluAsnCysProlLysAspAs 751  
Db 2459 CCACAGCTGCACAGAAGGCCAGATCTCCATTGCCAAGTATGAGAGTTGCCCCCAAGATAA 2518  
Qy 751 nProMetTrrTyrCysAsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnC 771  
Db 2519 CCCCCTGTACTACTGCAATAAGAAAACCCAGCTGCAGAGAGCTGTGCCCTTAGACCAAGACTG 2578  
Qy 771 sGlnTrrPLeuProArgAsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGI 791  
Db 2579 CCAGTGGGAGCCCGGATCAAGATGCAATGCCCTGCGGNAATATCTGTGGCAATGG 2638  
Qy 791 yTrrHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAs 811  
Db 2639 CTGGCATTTGGTTGGAACTCGTGTCTGAAATCACTACTGCTAAGAGAGAAATTATGACA 2698  
Qy 811 nAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrLysLy 831  
Db 2699 TGCTAAATTTGCTGTAGAACCAACATGCCCTTTTGGCTTCCCTTCCATCCAGAGAA 2758  
Qy 831 sValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLe 851  
Db 2759 GGTGGAGTTTGTCTTAAGCAGCTTCGATTAATGCAATCATCTCAAGATATGTCCAGCT 2818  
Qy 851 uThrLeuThrProTrrPValGlyLeuArgLysIleAsnValSerTyrTrrCysTrrPLeuAs 871  
Db 2819 CACTCTGATCTCATGGTGGTGTCTTGGAGAGATCAATGTGTCTTACTGTCTGTCTGGAGGA 2878  
Qy 871 pMetSerProPheThrAsnSerLeuLeuGlnTrrPMetProSerGluProSerAspAlaGI 891  
Db 2879 TATGTCTCAATTCACAAATAGTTTCTGCAGTGGATGCCATCTGAGCCAGTGTGCTGG 2938

Qy 891 yPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAs 911  
Db 2939 CTTCTGTGGGATCTTGTTCAGAGCTAGTACTCGGGATTAAAGGCTCAACCTGCATCAA 2998  
Qy 911 nProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgTh 931  
Db 2999 CCCTCTCAATGGGAGCCTCTGTGAAGGCTTGAACCCACAGTGCACAGAGTGCAGGAC 3058  
Qy 931 rProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTr 951  
Db 3059 ACCATGTGCCCTCGGACAGCGTGTGGCAGTGCATAGCAGCAGCTCGGAGTGCATGTG 3118  
Qy 951 pCysSerAsnMetLysGlnCysValAspSerAsnAlaTrrValAlaSerPhePropheGI 971  
Db 3119 GTGCAGTAACATGAAGCAGTGTGTGGATCTCCATGCTACGTGGCTCTCTCTCTTTGG 3178  
Qy 971 yGlnCysMetGlnTrrTrrThrMetSerThrCysProProGluAsnCysSerGlyTrrCy 991  
Db 3179 CCAGTGTATGGATGGTATACGATGAGCAGCTGCCACCTGAAATTTGCTTGGTACTG 3238  
Qy 991 sThrCysSerHisCysLeuGlnProGlyCysGlyTrrCysThrAspProSerAsnTh 1011  
Db 3239 TACCTGCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTGGTGTACTGATCCTAGCAATAC 3298  
Qy 1011 rGlyLysGlyLysCysIleGluGlySerTrrLysGlyProValLysMetProSerGlnAl 1031  
Db 3299 TGGGAAAGGAAATGTATTGAGGCGAGCTATAAGGACCTGTGAAGATGCGCTCACAGGC 3358  
Qy 1031 aProThrGlyAsnPheTrrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSe 1051  
Db 3359 CTCTGCAGGAAATCTGTATCCACAGCCCTTCTGAACTCCAGCATGTGTCTAGAGGACAG 3418  
Qy 1051 rArgTyrAsnTrrSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCy 1071  
Db 3419 CAGATACAACCTGGTCTTTCATTCTCAGTGTCCAGCTTCCAGTGCACACGACACGCAATG 3478  
Qy 1071 sIleAsnGlnSerIleCysGlyLysCysGluAsnLeuThrThrGlyLysHisCysGluTh 1091  
Db 3479 CATCAACACAGATATCTGTGAAGTGTGAGGACCTGACCACCGGACAGCAGCTCGGAGAC 3538  
Qy 1091 rCysIleSerGlyPheTrrGlyAspProThrAsnGlyLysCysGlnProCysLysCy 1111  
Db 3539 CTGCATATCTGGCTTCTATGCTGACCGCACTAATGAGGCAATGTTCAGCATGCAAGTG 3598  
Qy 1111 sAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGI 1131  
Db 3599 CAATGGGCACGCACTACTGTGCACACCAACACACCGGCAAGTGTCTGTACCCCAAGG 3658  
Qy 1131 yValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgTrrGlnGlyAsnProLe 1151  
Db 3659 TGTCAAGGGGACGAGTGCAGCTATGTGAGGTAGAAAATCGATACCAAGGAAACCTCT 3718  
Qy 1151 uArgGlyThrCysTrrTrrThrLeuLeuIleAspTrrGlnPheThrPheSerLeuSerGI 1171  
Db 3719 CAAAGGAACATGCTACTATACCTTCTCATTTGATCATGTTCCACCTTTAGCCCTGTCCCA 3778  
Qy 1171 nGluAspAspArgTrrTrrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnAr 1191  
Db 3779 GGAAGACGACCGCTACTACACAGCCATCACTTTTGGTGTACTCTCTGATGAAACAAACAG 3838  
Qy 1191 gAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrrAlaAl 1211  
Db 3839 GGATTTGGCATGTTTCATCAATGCTTCCAAAAAATCTCAACCTCAACATCACTCGGGCCAC 3898  
Qy 1211 aSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnI 1231  
Db 3899 CAGCTTCCCAACCGGAACCCAGACTGGAGAAGAGGTGCTGTGTTTCAAAAAACCAACAT 3958  
Qy 1231 eLysGluTrrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnI 1251  
Db 3959 CAAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGATTTTCGCAACCATCAACAT 4018

```
QY 1251 eThrPhePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSe 1271
Db 4019 CACTTCTTTGTTATGTAGTAATTTCACTTGGGCCCATCAAAATTCAGATTGCTTCTC 4078
QY 1271 rGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLe 1291
Db 4079 CCAGCACACCAACTTCATGGAGCTGCTGACAGTCTTCTCGTACTTCTTCTTCTTCTTCTTCT 4138
QY 1291 uSerLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTTPAlaSerAr 1311
Db 4139 CTCGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4198
QY 1311 gArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgPropheAlaSerVa 1331
Db 4199 GCGGAGAGAGCAACTTCTTCGGGAGATGCAACAGATGGCCAGCCGCCCTTTCCTCTCTGT 4258
QY 1331 lAsnValAlaLeuGluThrAspGluGluProAspIleLysGlySerIleLysTh 1351
Db 4259 AAACGTTGCTTGGAAACAGATGAGAGCTTCTGATCTTATTTGGGGGAGATATAAGAC 4318
QY 1351 rValProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerVa 1371
Db 4319 TGTTCCCAACCCATTGCACTGAGCCGCTGTTTGGCAACAAAGCCGCTGCTCTCTGT 4378
QY 1371 lPheValArgLeuProArgGlyLeuGlyIleProProProProGlyGlnSerGlyLeuAl 1391
Db 4379 GTTGTGAGGCTCCCTCGAGGCTGGTGGCATCCCTCCTCTGGGCGAGTCAGGTCTTTC 4438
QY 1391 aValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSe 1411
Db 4439 TGTGGCCAGCCCTCGTGGATCTTCTCAGCAGATGCGGATAGTGTACAGGAAATC 4498
QY 1411 rGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429
Db 4499 AGGAGCGGTGAGAAACCGGAAGCAGCAGCCCTCGCAGCCCTGGGAGCTGCATC 4553

RESULT 3
US-09-245-041-16
; Sequence 16, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16
; LENGTH: 4072
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-16

Alignment Scores:
Pred. No.: 0 Length: 4072
Score: 6870.50 Matches: 1240
Percent Similarity: 97.41% Conservative: 3
Best Local Similarity: 97.18% Mismatches: 20
Query Match: 86.80% Indels: 13
DB: 3 Gaps: 3

US-09-787-097-12 (1-1429) x US-09-245-041-16 (1-4072)
QY 2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuAlaArgLeuAlaAlaThrAlaAla 21
Db 10 GTGGCCGAGCGCGGCACTGAGGCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 69
```

```
QY 22 LeuAlaGlyArgSerGlyGlyProHisTrpAspValThrArgAlaGlyArgPro 41
Db 70 CTCGGGGCAGGAGCGCGGGCCGACCCGCTGCACCGCAGACAGGGGCGCTGGAGGCG 129
QY 42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56
Db 130 GGAACCGCGCGCGGTGTCTCCCGCGGGTGTCTCGCGGGCGGTGTCTCCCGCGCGCG 189
QY 57 LeuArgProArgLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeu 76
Db 190 CTGCTGCGG---CTGCTCTTTTCGCTGCTGCTGCTGCG----- 225
QY 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaValSerGlySerAlaAla 96
Db 226 CTGCGCGCGGAGCGCGCGCTGCGCGCGCGCGCGCGCTGCGCGCGCGCGCGCG 285
QY 97 GluAlaLysGluCysAspArgProCysValAsnGlyGlyArgCysAsnProGlyThrGly 116
Db 286 GAGCCCAAGGAATGTGACCGCGCGTGTCAACCGCGGTGCTGCAACCTCTGGCAGCG 345
QY 117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136
Db 346 CAGTGGGTCTGCGCGCGCGCTGGGTGGCGAGCAATGCCAGCACTGCGGGGCGCGCTTC 405
QY 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysThr 156
Db 406 AGACTAATCGATCTTCTGGGTGTGTGACAGATGGACCTGGAAATATATAATACAAACG 465
QY 157 LysCysThrTrpLeuIleGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176
Db 466 AAGTGCACGTGGCTCATTTGAGGACAGCAATAGATAATAGACTTCTGTTCAATCAT 525
QY 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196
Db 526 TTTGCTACAGAGTGTAGTTGGGACCATTTATATATATATATATATATATATATATATATGA 585
QY 197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216
Db 586 CCGCTAGTGTGCTGCAATTTAGTGGCTCATTTGTTCTTGAGAGAGATGCGCAATGAGACTGTC 645
QY 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaTyr 236
Db 646 CTTGAGGTTTGTGCCACATCAGGTTATGCTTCTGCTGCTATTTTATGATGCTGCTTAT 705
QY 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256
Db 706 AATTTGACTGGATTTAATATTACTTACAGTTTGTATATGTTGCCAATTAAGTCTCAGGC 765
QY 257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276
Db 766 CGAGGAGAGTGAAGATCAGTAATAGCAGCATGTTGTAATGTGAATGTGAATGTGAATGAAC 825
QY 277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296
Db 826 TGGAAAGGTGAAGCATGTGACATTCCTCTCATCTGACAGACAACTGTGCTTCTCTCATCGA 885
QY 297 GlyTleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316
Db 886 GGCATCTGCATTTCAAGTGTGATGTCAGAGGATGCTCTCTGCTTCTCAGACTGGCAGGCTCCT 945
QY 317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsn 336
Db 946 GGATGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1005
QY 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356
Db 1006 TTAAGCTCCCGAGGATCTCATAAAGCTGTGCTCAATGGAACATTTATGTGGTGTGT 1065
QY 357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376
Db 1066 GGAGGATATATGTTCAACCACTCAGATTTAATGATGTTCTAGCGTATGACCTTGCTTCT 1125
```

QY 377 ArgGluTyrLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSer 396  
DB 1126 AGGAGTGGCTTCCACTAAACCGTCTCTGTGAACATGTGGTGTGTAGATATGGCTACTCT 1185  
QY 397 LeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsn 416  
DB 1186 TTGGCATATATACAGGATAAAATTTACATGTATGAGGAGAAAATTTGATTCACCTGGGAAT 1245  
QY 417 ValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTyrValLeuLeuThrPro 436  
DB 1246 GTGACCATGTAGTTGAGAGTTTTCACATTCATAATGAGTGTATGGGTGTGTGACCCCT 1305  
QY 437 LysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsn 456  
DB 1306 AAGGCAAAAGGACGATGTCAGTGTGGGCACTCTGCACACATTTGTACACTGAAGAAT 1365  
QY 457 GlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsn 476  
DB 1366 GCGCGAGTGGTCACTGTGTCATCTTTGGTCACTGCGCTCTCTATGGGATATATAGCAAT 1425  
QY 477 ValGlnGluTyrAspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeu 496  
DB 1426 GTGCAGGAATATGATTGGATNAAGACACATGGAGTATATTACACACCGAGGTGCCCTT 1485  
QY 497 ValGlnGlyTyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrVal 516  
DB 1486 GTGCAAGGGGGTTACGGCATAGCAGTGTTTACGACCATAGACACGAGCGCCCTATACGTT 1545  
QY 517 HisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAspLeuTyrArg 536  
DB 1546 CATGTGTGCTACAGGCTTTGATGCCAATAGTACCGCTTGCAGATGATCTCTACCGA 1605  
QY 537 TyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPheArgTyrLeu 556  
DB 1606 TATGATGTGGATACCCAGATGGGACCATCTTAAAGGACAGCGCATTTTCCGTACTTGT 1665  
QY 557 HisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAsp 576  
DB 1666 CACACAGCTGTATAGTAGTGGAACCATGTGCTGGTGTGGGGGAAACACACAAATGAC 1725  
QY 577 ThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAla 596  
DB 1726 ACATCTATGAGCCATGGCGCAATGCTTCTCTTCAGATTTTCATGGCCCTATGACATTGCC 1785  
QY 597 CysAspArgTrpSerValLeuProArgProAspLeuHisHisAspValAsnArgPheGly 616  
DB 1786 TGTGACCGCTGTGTAGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTTTGGC 1845  
QY 617 HisSerAlaValLeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeu 636  
DB 1846 CATTCAGCAGTCTTACACACAGCACCATGATGTGTTCCGGTGGTTTCAATAGTCTCCTC 1905  
QY 637 LeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAla 656  
DB 1906 CTCAGCGACATCTGTGATTCACCTCGGAACAGTGTGATGGCATCGGAGTGAAGCGCT 1965  
QY 657 CysLeuAlaAlaGlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIle 676  
DB 1966 TGTTAGCAGAGGACCTGGTATTCCGGTGTGTGGAAACACAGGTCTGCTCAGTGTATC 2025  
QY 677 SerTrpAlaLeuAlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLys 696  
DB 2026 TCGTGGCGCTGGCAACTGATGAACAAGAGAAAGTTAAATCAGATGTTTTCCTCAA 2085  
QY 697 ArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsn 716  
DB 2086 AGAACTCTTGACCATGTGACAGATGTACCAGCACACAGATTTGTACAGTGCACAGCCCAAC 2145  
QY 717 ThrAsnAspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGlu 736  
DB 2146 ACCAATGACTGTCATGTGTGCANTGACCATGTGTGTCAGGAAACACAGCTGTCTCAGAA 2205  
QY 737 GlyGlnIleSerIlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCys 756

DB 2206 GGCGAGATCTCCATTTTATAGTATGAGATTTGCCCAAGGATAACCTATGACTACTGT 2265  
QY 757 AsnLysLysThrSerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArg 776  
DB 2266 AACAAAGAACCAAGCTGCGAGAGCTGTGCCCTGGACAGAACTGCCAGTGGAGCCCGG 2325  
QY 777 AsnGlnGluCysIleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGly 796  
DB 2326 AATCAGGAGTGCATTTGCCCTGCCGGAATAATCTCTGTGGCATTTGGCTGGCATTTGGTTGA 2385  
QY 797 AsnSerCysLeuLysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCys 816  
DB 2386 AACTCATGTTTGAATAATTAATACTGCAAGGAGAAATATGACAATGCTAAATTTGTTCTGT 2445  
QY 817 ArgAsnHisAsnAlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeu 836  
DB 2446 AGAACACACAAATGCCCTTTTGGCTTCTCTTACAAACCCAGAGAGGTAGAAATTTGTCCTT 2505  
QY 837 LysGlnLeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrp 856  
DB 2506 AAGCAGCTGCGAATAATGCAATCATCTCAGAGCATGTCCAAAGCTCACTTAACCCCATGG 2565  
QY 857 ValGlyLeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThr 876  
DB 2566 GTCGGCTTCCGAGATCAATGTGCTCTACTGTGTGGGAAGATATGTCCCATTTTACA 2625  
QY 877 AsnSerLeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeu 896  
DB 2626 AATAGTTTACTACAGTGGATGCCCTCTGAGCCAGTGTGATGCTGATCTCTGTGGAATTTTA 2685  
QY 897 SerGluProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySer 916  
DB 2686 TCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGT 2745  
QY 917 ValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArg 936  
DB 2746 GTCTGTGAAGGCTGCCAAACCAACAGTGTACAGTGTCCGACACCATGTGCTTGAGG 2805  
QY 937 ThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLys 956  
DB 2806 ACAGCATGTGGAGATTGCACCGCGCAGCTCTGAGTGCATGTGTGTGACCAACATGAAG 2865  
QY 957 GlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrp 976  
DB 2866 CAGTGTGTGACTCCAAATGCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGG 2925  
QY 977 TyrThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCys 996  
DB 2926 TATACGATGACACCTGCCCTGCTGAAATTTGTTTACGGCTACTGTACCTGTACTCATTCG 2985  
QY 997 LeuGluGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCys 1016  
DB 2986 TTGGAGCAACAGGCTGTGGCTGTGTACTGTATCCAGCAATATCTGGCAAGGGAATGC 3045  
QY 1017 IleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPhe 1036  
DB 3046 ATAGAGGTTTCTTATPAGGACCAAGTGAAGTGCCTTCCGACGCCCTTACAGAAATTTTC 3105  
QY 1037 TyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSer 1056  
DB 3106 TATCCACAGCCCTGTCTCAATTCAGCATGTGTCTAGAGACAGCAGATACAACTGGTCT 3165  
QY 1057 PheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIle 1076  
DB 3166 TTCAATCAGTGTCCAGCTTGCCTGCAATGCAACGGCCACAGTAAATGCAATCAATCAGACATC 3225  
QY 1077 CysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPhe 1096  
DB 3226 TGTGAGAGTGTGAGAACCTTGACCAAGGACAGCACTGGGAGACCTGCATATCTGCTTC 3285  
QY 1097 TyrGlyAspProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSer 1116

Db 3286 TAGGTGATCCACCAATGAGGAGAAATGTCAGCCATGCAAGTGCATGAGGACGCGTCT 3345  
Qy 1117 LeuCyAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGlu 1136  
Db 3346 CTGTGCAACACCAACACAGGCGAAGTCTTCTGCACCAACCAAGGCGTCAAGGGGACGAG 3405  
Qy 1137 CysGlnLeuGlyValGluAsnArgTyrglnGlyAsnProLeuArgGlyThrCysTyr 1156  
Db 3406 TGCCAGCTATGAGGTAGAAATGATACCAAGGAAACCCCTCTCAGAGGAACATGTTAT 3465  
Qy 1157 TyrThrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1176  
Db 3466 TATACCTCTTATTGACTATCAGTTTCACTTACCTTATGCTATATCCAGAGGATGATCGCTAT 3525  
Qy 1177 TyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPhe 1196  
Db 3526 TACACAGCTATCAATTTTGTGCTACTCTCTGACGAAACCAACAGGGATTTGACATGTTT 3585  
Qy 1197 IleAsnAlaSerLysAsnPheAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 1216  
Db 3586 ATCAATGCCCTCCAGAAATTTCACTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGA 3645  
Qy 1217 ThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrllysasp 1236  
Db 3646 ACCAGGCTGGAGAGAGATGCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGAT 3705  
Qy 1237 SerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyr 1256  
Db 3706 AGTTTCTCTAAAGAGAGTTTGAATTTTCGAACCAACCCCAATATACCTTCTTTGTTTAT 3765  
Qy 1257 ValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGln 1272  
Db 3766 GTCAGTAATTTTCCCTGGCCCATCAAAATTCAGGTGCAAACTGAACAA 3813

## RESULT 4

US-09-245-041-18  
; Sequence 18, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-18

Alignment Scores:  
Pred. No.: 0 Length: 2625  
Score: 4136.50 Matches: 759  
Percent Similarity: 95.48% Conservative: 2  
Best Local Similarity: 95.23% Mismatches: 22  
Query Match: 52.26% Indels: 14  
DB: 3 Gaps: 3

US-09-787-097-12 (1-1429) x US-09-245-041-18 (1-2625)

Qy 2 ValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 21  
Db 10 GTGGCGCAGCGCGCGCAACTGAGGCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 69  
Qy 22 LeuAlaGlyArgSerGlyGlyProHisTrpAspValThrArgAlaGlyArgPro 41

Db 70 CTCGGGCGCAGGAGCGCGCGCGCGCGCACCGACCCCTGCACCGCGACAGGGGCGCTGGAGCGCG 129  
Qy 42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56  
Db 130 GGACCGCGCGCGCGCGCGCTGCTCTCCGCGCGGCTGCTCGCGCGCGCTGCGCGCGCGCGCG 189  
Qy 57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 76  
Db 190 CTGCTGCGG---CTGCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 225  
Qy 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 96  
Db 226 CTGCGCGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 285  
Qy 97 GluAlaLysGluCysAspArgProCysValAsnGlnGlyArgCysAsnProGlyThrGly 116  
Db 286 GAGGCCAAGGAATGTACCGCGCGTGTCAACGCGCGGCTGCTGCAACCCCTGGCAACCTGGCACCGCG 345  
Qy 117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136  
Db 346 CAGTGGCTCTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTTC 405  
Qy 137 ArgLeuThrGlySerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156  
Db 406 AGACTAATCTGATCTTCTGGGTTTGTGACAGATGGACCTGGAAATTTATAATACAAACG 465  
Qy 157 LysCysThrTrpLeuLeuGluGlyGlnProAsnArgIleMetArgLeuArgPheAsnHis 176  
Db 466 AGTGCAGTGGCTCAITGAGGACGACCCCAATAGAAATTTATGAGACTTGGTTTCATCAT 525  
Qy 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196  
Db 526 TTTGTACAGAGTGTAGTTGGGACCAATTTATATGTTTATGATGGGACTCAATTTATGCA 585  
Qy 197 ProLeuValAlaAlaPheSerGlyLeuLeuValProGluArgAspGlyAsnGluThrVal 216  
Db 586 CCCTAGTGTCTGATTTAGTGGCCCTCATTTGTTTCTGAGAGAGATGGCAATGAGACTGTC 645  
Qy 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaLafyr 236  
Db 646 CTTGAGGTTGTGCCACATCAGGTTATGCTTCTGCTGCAATTTTCTGATGCTGCTTAT 705  
Qy 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256  
Db 706 AATTTGACTGGATTTAATATTTACTTACGTTTGTATGTTGTTGTTGTTGTTGTTGTTGTTGTT 765  
Qy 257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276  
Db 766 CGAGGAGAGTGAAGATCAGTATAGCAGGATCTGTTGATGATGATGATGATGATGATGATGAT 825  
Qy 277 TrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArg 296  
Db 826 TGGAAAGGTGAAGCATGTGACATTTCTCTACGACACACTGTTGTTTCTCTCATCGA 885  
Qy 297 GlyTleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPro 316  
Db 886 GGCATCTGCAATTCAGTGTATGTCAGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCT 945  
Qy 317 GlyCysSerValProValProAlaAsnGlnSerPheTrpThrArgGluLysSerAsn 336  
Db 946 GGATGTTTCTGTTTACCAGCTAAACAGTCACTTTTGGACTCGAGAGGAATATTTCTAAC 1005  
Qy 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpValVal 356  
Db 1006 TTAAGCTCCCCAGAGCATCTCATTAAGCTGCTGCTCAATGGAACCATTTATGTTGGTGT 1065  
Qy 357 GlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSer 376  
Db 1066 GGAGATATATGTTTCAACCACTCAGATTTATAACATGTTTCTAGCGTATGACCTTGTCTTCT 1125  
Qy 377 ArgGluTrpLeuProLeuAsnArgSerValAsnValValValValValValValValValVal 396







Db 421 TGTGCTTGGATTCCTGTTGGGAATTTTATTCAGGAACCCAGTTACTTCGGGACCTGAAGG 480  
Qy 906 laAlaThrCysIle-AsnProLeu-AsnGlySerValCysGluArgProAlaAsnHis 925  
Db 481 CTGCAACCTGCTTCAACCCACTYAAATGGTAGTGTCTGTGAAGGCTGCAACACACAG 540  
Qy 925 rAlaIys-GlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerG 945  
Db 541 TGTAAAGCAGTGGCCGACACCATGTGCTTGGACACAGCATGTGGAGATTGCACACGG 600  
Qy 945 lYSerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTrp 965  
Db 601 GCAGCTCTCAGTGCATGTGTCGACCAACATGAAGCAGTGTGGACTCCAATGCTATG 660  
Qy 965 aAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProG 985  
Db 661 TGGCCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTTATACGATGAGCACCTGCCCCCTG 720  
Qy 985 luAsnCysSerGlyTrpCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpC 1005  
Db 721 AAAATTGTTACGGCTACTGTACTCTGTAGTCATCTGCTTGGAGCAACACAGCTGTGGCTGGT 780  
Qy 1005 yThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTrpLysGlyProV 1025  
Db 781 GTACTCATCCAGCAATACTGGCAAGGGAATGATAGAGGGTTCCTATAAGGACACAG 840  
Qy 1025 aIlyMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerS 1045  
Db 841 TGAAGATGCTTGGCAAGCCCTTACAGGAAATTTCTTCCAGCCCTGTCTCAATTCCA 900  
Qy 1045 exMetCysLeuGluAspSerArgTrpAsnTrpSerPheIleHisCysProAlaCysGlnC 1065  
Db 901 GCATGTGTCTAGAGGACACAGATCAACTGCTGCTTTTCACTCACTGTCCAGCTTGCCAA 960  
Qy 1065 yAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrT 1085  
Db 961 GCAACGGCCACAGTAAATGATCAATCAATCAGAGATCTGTGAGAAAGTGTGAGAACCTGACCA 1020  
Qy 1085 hrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyL 1105  
Db 1021 CAGGCAAGCACTGCGAGACTGCTGATATCTGGGTTTCTACGGTATCCCAATCGAGGGA 1080  
Qy 1105 yCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysC 1125  
Db 1081 AATGTCAAGCATGCAAGTCAATGGCAGCGCTCTGTGCAACACCAACACGGGCAAGT 1140  
Qy 1125 yPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnA 1145  
Db 1141 GCTTCTGCAACCAACAGGGCGCTCAAGGGGAGCAGTGCAGCATGTGAGGTAGAAATC 1200  
Qy 1145 rGTrpGlnGlyAsnProLeuArgGlyThrCysTyrThrLeuLeuIleAspTrpGlnP 1165  
Db 1201 GATACCAAGAAACCTCTCAGAGGAACATGTTATTACTCTCTTATTGACTATCAGT 1260  
Qy 1165 heThrPheSerLeuSerGlnGluAspAspArgTyrThrAlaIleAsnPheValAlaT 1185  
Db 1261 TCACCTTTTATCTATCCAGGAAGATGATCGGTATTATACACAGCATCAATTTTGGGTA 1320  
Qy 1185 hrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnL 1205  
Db 1321 CTCTGACGAACAAACAGGGATTGGACATGTTTCATCAATGCTCCCAAGAAATTCACAC 1380  
Qy 1205 euAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProV 1225  
Db 1381 TCAACATCACTGGGCTGGCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAGATGCGCTG 1440  
Qy 1225 aIValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspP 1245  
Db 1441 TTGTTTCAAAACCAACATTAAGGATGACAAAGATAGTTTCTTAATGAGAAGTTTGATT 1500  
Qy 1245 heArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleL 1265  
Db 1501 TTGCAACCAACCAATATACCTTTCTTTGTTTATGTGATTAATTCACCTGGCCCATCA 1560

Qy 1265 ysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValT 1285  
Db 1561 AAATTTCAGATTGCTTCTCTCAGCACAGCAATTTTATGACCTGGTACAGTTCTTCGTGA 1620  
Qy 1285 hrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrrpLysIleLysG 1305  
Db 1621 CTTTCTTCAGTTGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1680  
Qy 1305 lnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaS 1325  
Db 1681 AAAGTTGTTGGGCTCCAGACGTAGAGACAACTTCTCGAGAGATGCAACAGATGGCCA 1740  
Qy 1325 exArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuI 1345  
Db 1741 GCCGTCTCTTGGCTTGTAAATGTCCCTTGGAAACAGATGAGGAGCTCTCTGATCTTA 1800  
Qy 1345 leGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnL 1365  
Db 1801 TTGGGGGAGTATAAAGACTGTTCCCAAAACCCATTGCACTGGAGCGCTGTTTGGCAACA 1860  
Qy 1365 ysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProP 1385  
Db 1861 AAGCCGCTGCTCTCTGTTTGTGAGGCTCCCTCGAGGCTTGGGTGGCATCCCTCTCTC 1920  
Qy 1385 roGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProI 1405  
Db 1921 CTGGGCACTCAGCTCTTGTGTGGCCAGCGCTGCTGTGACATTTCTCAGCAGATGCCGA 1980  
Qy 1405 leValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnP 1425  
Db 1981 TAGTGTACAGGAGAAAGTCAGGAGCGCTGAGAAACCGGAAGCAGAGCCCTGCAAGC 2040  
Qy 1425 roGlyThrCysIle 1429  
Db 2041 CTGGGACCTGATC 2054

RESULT 6  
US-09-245-041-8  
; Sequence 8, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-8

Alignment Scores:  
Pred. No.: 1,71e-301 Length: 2419  
Score: 3435.00 Matches: 642  
Percent Similarity: 84.37% Conservative: 22  
Best Local Similarity: 81.58% Mismatches: 44  
Query Match: 43.40% Indels: 79  
DB: 3 Gaps: 4

US-09-787-097-12 (1-1429) x US-09-245-041-8 (1-2419)

Qy 2 ValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaAla 21

83 CTGGCGGCGCGCGCGCTGAGCGGGGCTGAGGGGAGCACGAGGACGACAGCG 142  
22 LeuAlaGlyArgSerGlyGlyProHisTrpAspValThrArgAlaGlyArgPro 41  
143 CTGGCGGCGAGGAGGCGAGGACCGACCTGACCGGACGAGGGGCTTGAGGCGG 202  
42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56  
203 GGACCGCGCGCGCGCTGTCTCCGCGGGGTGTCTCGCGGGCGCTGCCCGCGCGCG 262  
57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeu 76  
263 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCCG----- 298  
77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaValSerGlySerAlaAlaAla 96  
299 CTGCCCCGGAGCGCGCGCTGCGGTGGCGGGCGGGTCTCGGCTCGCGCGAGCC 358  
97 GluAlaLysGluCysAspArgProCysValLanGlyGlyArgCysAsnProGlyThrGly 116  
359 GAGGCGCAAGGAATGTGACCGGCGGTGTCTCAACCGCGCGCGGTGCAACCTGGCACCGGC 418  
117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyClyArgPhe 136  
419 CAGTGGCTGTGCCCGCCAGCGGTGGTGGCGAGCAATGCCAGCACTGGCGGGCGCGCTTC 478  
136 ----- 136  
479 AGGACATCTGCTCTCAGCGCTATAATCACAGCTGTTTCGGAAGGTGAGGCTGGAGGAACAGT 538  
136 ----- 136  
539 TCGAGCGAAGCTTCGGCTACAGANTAGTTCAGAGTAACCTGGGGCAACTTGGGCTTGT 598  
136 ----- 136  
599 CTCCAAAACCAAAATGAGCGAAAAGAGCAAGCTAGAGTCTTTTGGGAAAAATTTAGCTG 658  
137 -----Arg-LeuThrGlySerSerGlyPheValThrAspGlyProGlyAs 151  
659 ACTAAATTTTTCACCGAGAACTAACTGGCTCTTCGGATTTGTAAACAGATGACCTGGGAA 718  
151 nTrpLysTrpLysTrpLysCysThrTripleuileGluGlyGlnProAsnArgilleMetAr 171  
719 TTATAAATATAGAGAGAGTGCACATGGCTCATTCGAGACAGCCAAATAGANTAGAG 778  
171 gLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeuTrpValTrpAspGly 191  
779 ACTTCGCTTCAACCATTTTGTCTACAGAAATGTAGCTGGGACCAATTTATATTTATGATGG 838  
191 yAspSerileTrpAlaProLeuValAlaAlaPheSerGlyLeuileValProGluArgAs 211  
839 GGACTCAATCTACGCACTCTCATTTGCTGCTGCTTTAGTGGCTTCATTTCTGTAAGAGAGA 898  
211 pGlyAsnGluThrValProGluValAlaAlaThrSerGlyTrpAlaLeuLeuHisPhePh 231  
899 TGGCAATGAGACGGCTCTCTGAGGTCACTGCTCACTTCACTTATGCTGCTGCTATTTT 958  
231 eSerAspAlaAlaTrpAsnLeuThrGlyPheAsnileThrTrpSerPheAspMetCysPr 251  
959 CAGTGATGCTGCTTATCTCACTGGATTTAATATCACTTCAATTTTGAATGTGTGCC 1018  
251 oAsnAsnCysSerGlyArgGlyGluCysLysileSerAsnSerSerGluThrValGluCys 271  
1019 GAATAAATGCTCAGCCCGGAGAGTGTAAAGCAGTAACAGCAGCGGCTGTGTAGTG 1078  
271 sGluCysSerGluAsnTrpLysGlyGluAlaCysAspPheProHisCysThrAspAsnCy 291  
1079 TGAATGCTCTGAAAACTGAAAGGGGAGTGTGTGACATTCCTCACTGTACAGACAACTG 1138  
291 sGlyPheProHisArgGlyLysCysAsnSerSerAspValArgGlyCysSerCysPheSe 311  
1139 TGGCTTTCTCCACCGAGGATCTGAATGCAAGCGATACAGAGGGTGTCTCTGCTTTCC 1198

311 rAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGlnSerPheTrpThrAr 331  
1199 TCACCTGGCAGGCTCTGGATGTTCAATCTCTGTCGACAGCTAAACAGCTCTTTTGGACTCG 1258  
331 gGluGluTrpSerAsnLeuLeuLeuProArgAlaSerHisLysAlaValValAsnGlyAs 351  
1259 AGAAGAAATATCTGATTTAAAGCTTCCAGAGGCTCTCTATAAAGCTGTGTGTCATGAA 1318  
351 nIleMetTrpValValGlyGlyTrpMetPheAsnHisSerAspTrpAsnMetValLeuAl 371  
1319 TATAATGTGGGTGTGTGGCGGATATATGTTCAACCAATTCAGATTACAGCATGGTTTAGC 1378  
371 aTrpAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValVa 391  
1379 GTATGACCTGACTCTAGGGAATGGCTCCACTAAACCACTCTGTGTGAACAGATGTGGTTGT 1438  
391 lArgTrpGlyHisSerLeuAlaLeuTrpLysAspLysileTrpMetTrpGlyLysile 411  
1439 ARGATATGGTCATCTTTGGCATTAATAAGGATAAAATCTACATGTATGAGGAGAAAT 1498  
411 eAspProThrGlyAsnValThrAsnGluLeuArgValPheHisileHisAsnGluSerTr 431  
1499 TGATTCACAGGGAACGTGACCAATGAGCTGAGAGTATTTTCAATTCATAATGAATCAIG 1558  
431 pValLeuLeuThrProLysAlaLysGluGlnTrpAlaValValGlyHisSerAlaHisil 451  
1559 GGTATTTTAACTCCGAAAGCTAAGGATCAAGTATGCACTGGTGGACACTCAGCACACAT 1618  
451 eValThrLeuLysAsnGlyArgValValMetLeuValilePheGlyHisCysProLeuTy 471  
1619 TGTTACACTGCTGCTGGCGCTGTGGTCATGTTGGTCACTTCCTCGGTCACTGCCACTCTA 1678  
471 rGlyTrpIleSerAsnValGlnGluTrpAspLeuAspLysAsnThrTrpSerileLeuHi 491  
1679 TGGATATTAAGGCTGTGTCAGCAATATGACTTGGAAAAAGAACACATGGAGTATATTACA 1738  
491 sThrGlnGlyAlaLeuValGlnGlyTrpGlyHisSerSerValTrpAspHisArgTh 511  
1739 TACTCAGGGTGTCTTGTGCAAGGGGTATGTCACAGTAGTGTATTATGATGACAGGAC 1798  
511 rArgAlaLeuTrpValHisGlyGlyTrpLysAlaPheSerAlaAsnLysTrpArgLeuAl 531  
1799 CAAGGCTCTGACTGCTCATGTGCTACAGGCTTTTCAGCGCCCAACAAATACCGGCTGC 1858  
531 aAspAspLeuTrpArgTrpAspValAspThrGlnMetTrpTrpIleLeuLysAspSerAr 551  
1859 AGATGACCTACAGATACGATGCTGATCTCAGATGTGAGTGGNACCATGCTGTGTGGAGG 1918  
551 gPhePheArgTrpLeuHisThrAlaValileValSerGlyThrMetLeuValPheGlyGly 571  
1919 ATTTTTCGTTTACTTGATACAGCTGTGATGATGATGATGAGTGGNACCATGCTGTGTGGAGG 1978  
571 yAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMe 591  
1979 GAACACACACAAATGACACTTCCATGAGCCAGCGGTGCGCAATGCTCTCTCGGACTTCAT 2038  
591 tAlaTrpAspIleAlaCysAspArgTrpSerValLeuProArgProAspLeuHisAs 611  
2039 GGCTTATGACATTCCTTGTGACCGATGTGCTGCTTCCAGACCTGAGCTCCATCATGA 2098  
611 pValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrMetTrpValPheGlyGly 631  
2099 TGTCACAGATTTGGCCATTCACAGCTTTGTGTACAAACAGACACCATGATGTGTGTGGCGG 2158  
631 yPheAsnSerLeuLeuSerAspIleLeuValPheThrSerGluGlnCysAspAlaHi 651  
2159 CTTCAACAGCGCTCTCTCAGTACGCTTGTGGTCTTTTACCTCGAGCAGTGGCATGCACA 2218  
651 sArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyLysileArgCysValTrpAsnThrGl 671  
2219 CCGCAGTGAAGCTGTGTGTGTGTCAGCAGGACCTGTGTATCCGGTGTCTGTGTGGACACACA 2278

QY 671 ySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnGluGluLysLeuLysE 691  
Db 2279 GTGCTTCGATGACCTCTCGGAGTTGGCACTGAAGCAAGCAGAAAGTTAAATC 2338  
QY 691 rGluCysPheSerLysArgThrLeuAspHisAspArgCysAspGlnHisThrAspCysTy 711  
Db 2339 AGAGTGTTTTCTAAAGAACCTTTGACCATGACAGATGTGACCGACGACACAGATTGTTA 2398  
QY 711 rSerCysThrAlaAsnThr 717  
Db 2399 CAGCTGCACAGCAATACC 2417  
RESULT 7  
US-09-245-041-10  
; Sequence 10, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; EARLIER FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 10  
; LENGTH: 1051  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-10  
Alignment Scores:  
Pred. No.: 1 52e-101 Length: 1051  
Score: 1226.50 Matches: 236  
Percent Similarity: 84.45% Conservatives: 3  
Best Local Similarity: 83.39% Mismatches: 31  
Query Match: 15.50% Indels: 13  
DB: 3 Gaps: 3  
US-09-787-097-12 (1-1429) x US-09-245-041-10 (1-1051)  
QY 2 ValAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAlaAla 21  
Db 217 GTGGCGCGCGCGCGCGCTGAGCGCGCGCTGAGCGGAGCAGCAGCAGCAGCG 276  
QY 22 LeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArgPro 41  
Db 277 CCTGCGCGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 336  
QY 42 GlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSer-----ProPro 56  
Db 337 GGACCGCGCGCGCGCGCTGCTCCCGCGGCTGCTCGCGCGCGCTGCGCGCGCG 396  
QY 57 LeuArgProArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 76  
Db 397 CTGCTGCCG---CTGCTCTTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 432  
QY 77 LeuProCysGluAlaGluAlaAlaAlaAlaAlaAlaAlaAlaValSerGlySerAlaAla 96  
Db 433 CTGCGCGCGGAGCGCGCGCGCTGCGGCGCGCGCGCGCTGCGGCTGCGCGCGCG 492  
QY 97 GluAlaLysGluCysAspArgProCysValAlaGlyArgCysAsnProGlyThrGly 116  
Db 493 GAGGCCAAGGAATGTACCGCGCGCTGTGCAACGCGCGCGCTGCAACCTGGACCGCG 552  
QY 117 GlnCysValCysProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPhe 136  
Db 553 CAGTGGCTGTGCCCGCGCGCTGGTGGCGGAGCAATGCCAGCACTTGGCGGCGCGCTTC 612

QY 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrlLysTyrlLysThr 156  
Db 613 AGACTAACTGGCTCTTCTGGATTTGTAAACAGATGGACCTGGGAATTAATAATAAGACG 672  
QY 157 LysCysThrTrpLeuLeuGluGlyGlnProAsnArgIleValArgLeuArgPheAsnHis 176  
Db 673 ARGTCACATGGCTCATTTGAAGGACACCAATAGAAATTAATGAGACTTCGCTTCAACCAT 732  
QY 177 PheAlaThrGluCysSerTrpAspHisLeuTyrlValTyrlAspGlyAspSerIleTyrlAla 196  
Db 733 TTTGCTACAGANTGTAGCTGGGACCAATTTATATGTTATGATGGGACTCAATCTACGCA 792  
QY 197 ProLeuValAlaAlaPheSerGlyLeuLeuValProGluArgAspGlyAsnGluThrVal 216  
Db 793 CCTCTGATTCGCTCTTATAGTGGCCTCATTTCTCGAAAGAGATGGCAATGAGCGCT 852  
QY 217 ProGluValValAlaThrSerGlyTyrlAlaLeuLeuHisPhePheSerAspAlaAlaTyrl 236  
Db 853 CCTGAGTCACTGCTCACTTCAGGTATGCACTGCTGCTCATTTTTCAGTCACTGCTGCTTAT 912  
QY 237 AsnLeuThrGlyPheAsnIleThrTyrlSerPheAspMetCysProAsnAsnCysSerGly 256  
Db 913 AATCTGACTGGATTTAATATCACTTACAATTTTGCATGCTCGCAATTAATTTGCTCAGCG 972  
QY 257 ArgGlyGluCysLysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsn 276  
Db 973 CGAGGAGAGTGTAAAGACGAGTAAACAGCAGCGCTGTTGAGTGTGATGTTCTGAAAC 1032  
QY 277 TrpLysGly 279  
Db 1033 TCGAAAGGG 1041  
RESULT 8  
US-09-245-041-5  
; Sequence 5, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; EARLIER FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 90050  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-245-041-5  
Alignment Scores:  
Pred. No.: 2 73e-56 Length: 90050  
Score: 766.00 Matches: 205  
Percent Similarity: 21.41% Conservatives: 2  
Best Local Similarity: 21.20% Mismatches: 6  
Query Match: 9.68% Indels: 755  
DB: 3 Gaps: 2  
US-09-787-097-12 (1-1429) x US-09-245-041-5 (1-90050)  
QY 849 SerIysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrlTrpCys 868  
Db 28996 TCCAGCTCACTGACTCCATGGGTGGTCTTCGAAAGATCAATGTGCTTACTGCTGC 29055  
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpValProSerGluProSer 888



QY	983	----	983
Db	31275	ACAGTGTTCACGACGATAAATCAGTTTACATCAAGGAGGACACAGGTCACAGAAATACC	31334
QY	983	----	983
Db	31335	TTTTCTCTTCAGGGTAAATATTATAAATTCAAATCTGTATAATGTTTCTACATCTTAAATCTA	31394
QY	983	----	983
Db	31395	CCAGTATGTAAGTGTCTTCTAGTAGAGGCCTCCACAGCTCCCTTTTTCATCCACACATCC	31454
QY	983	----	983
Db	31455	TGATATAAAGGTTGGAAGTCCCTGTTATATATTATGAAAAATGCGGGCCCTTTAA	31514
QY	983	----	983
Db	31515	ATTATTTCAGTTCATAATCACTATAGGTACTATTTTAAATTCATGGAAGTTAAATCAT	31574
QY	983	----	983
Db	31575	CTGTTAAAAAGAAAGTAATAACAGTAAATCAAATCTTGTATAGTGAATTACAGATTG	31634
QY	984	-----ProGluAsnCysSerGlyTYrCysThrCysSerHi	995
Db	31635	GATTGTTTTGCCTTGTTTTTAATAGCTGAAATTTGCTGGCTACTGTACTCTCAGCCA	31694
QY	995	sCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGly	1015
Db	31695	TTGCTTGGAGCAGCCAGGCTGGTGGTGTACTGATCCTAGCAATCTCGGAAGGANA	31754
QY	1015	sCysIleGluGlySerTYrLysGlyProValLysMetProSerGlnAlaIap-oThrGlyAs	1035
Db	31755	ATGATTTAGGGCAGCTATAAAGGACCTGTGAAGATGCCGTACACAGGCCTCTGCAGGAAA	31814
QY	1035	nPheTYrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerA-oGlyAsnTr	1055
Db	31815	TGTGTATCACAGCCCCCTCTGAACTCCAGCATGTCTTAGAGGACAGCATACACTG	31874
QY	1055	pSerPheIleHisCysPro	1061
Db	31875	GTCTTTTCATTCACTGTCCA	31893

```

RESULT 9
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1916
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(415)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

```

Alignment Scores:		
Pred. No.:	1.11e-43	Length: 415
Score:	581.50	Matches: 105
Percent Similarity:	84.56%	Conservative: 10
Best Local Similarity:	77.21%	Mismatches: 20

```

Query Match: 7.35%
DB: 4
Indels: 1
Gaps: 1

US-09-787-097-12 (1-1429) x US-09-833-381-1916 (1-415)

Qy 141 SerSerGlyPheValThrAspGlyProGlyAsnTyLysTyLysThrLysCysThrTrp 160
   : : : : :
Db 4 GCGTCGGGATATTAAACAGATGSCCCAAATAACTATAATAATAAACTAAATGTACTGG 63

Qy 161 LeuLleGluGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180
   : : : : :
Db 64 CTCATTGAAGGCTATCCAAATGCGAGTGTTAAGATTAAAGATTCATTITGTCACGAA 123

Qy 181 CysSerTrpAspHisLeuTyrValTyAspGlyAspSerIleTyrAlaProLeuValAla 200
   : : : : :
Db 124 TGTAGCTGGGATCATATGATGTATTTATGATGGAGATTCAAATATATGACCTTTAATAGCT 183

Qy 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220
   : : : : :
Db 184 GTACTTAGTGGTTTGATAGTCCCTGAAATAGGGGCATGAAACTGTGCTGAGTTGTT 243

Qy 221 AlaThrSerGlyTyrrAlaLeuLeuHisPhePheSerAspAlaAlaTyrrAsnLeuThrGly 240
   : : : : :
Db 244 ACTACATCTGGCTATGCACCTGTACATTTTTTTTAGTGATGCTGCGTATATCTTAACCTGGT 303

Qy 241 PheAsnIleThrTyrrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260
   : : : : :
Db 304 TTCACATTTTCTATTTCATCATCAATCTTGTCTCAACAATGCTCTGCTCATGSGAAGTGT 363

Qy 261 LysIleSerAsnSer---SerGluThrValGluCysGluCysSerGlu 275
   : : : : :
Db 364 ACAACTAGTCTCTCTGTCACAGCTCAAGTATATTTGTGAATGTCATAAA 411

```

RESULT 10

US-09-245-041-3/c

; Sequence 3, Application US/09245041

; Patent No. 6274339

; GENERAL INFORMATION:

; APPLICANT: Moore, K.

; APPLICANT: Nagle, D.

; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT

; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY

; FILE REFERENCE: 7853-136

; CURRENT APPLICATION NUMBER: US/09/245,041

; CURRENT FILING DATE: 1999-02-05

; EARLIER APPLICATION NUMBER: 60/093,630

; EARLIER FILING DATE: 1998-07-21

; EARLIER APPLICATION NUMBER: 60/104,978

; EARLIER FILING DATE: 1998-10-20

; NUMBER OF SEQ ID NOS: 131

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 3

; LENGTH: 17056

; TYPE: DNA

; ORGANISM: Mus musculus

US-09-245-041-3

Alignment Scores:		
Pred. No.:	3 1e-29	Length: 17056
Score:	454.00	Matches: 155
Percent Similarity:	16.81%	Conservative: 5
Best Local Similarity:	16.28%	Mismatches: 10
Query Match:	5.74%	Indels: 783
DB:	3	Gaps: 2

US-09-787-097-12 (1-1429) x US-09-245-041-3 (1-17056)

Qy	202	PheSerGlyLeuIleValProGluAtrGaspGlyAsnGluThrValProGluValValala	221
Db	5221	TTTCAGTGGCCTCATTGTTCTCGAAAGAGATGGCAATGACGGCTCCTGAGGTCACTGTC	5162
Qy	222	ThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaIatYrAsnLeuThrGlyPhe	241

5161	Db	ACTTCAGGTTATGCACCTGCTGCATTTTTTCAGTCATGCTGCTTTATAATCTGACCTGGATTT	5102
242	Qy	AsaIleThrTyr 	245
5101	Db	AATATCACTTTACAAGTAAGATATTAGATTTAAATTTTGTGATACTACTCAAATGACTTA	5042
245	Qy	-----	245
5041	Db	ATAACACTAATATATTGGAGGGAGGCTTCTAGCTAGTCTCAGCTCCTTGTAAATTTTAAAC	4982
245	Qy	-----	245
4981	Db	ATGGATTTTCTTCTTCTCTTTTTTTTTTTTTTTTTCATAATTGAAGAAACAGGAGCTAAAG	4922
245	Qy	-----	245
4921	Db	GTTAAACAAATATCAGCCCCACTGGTTGGGACCAAGTTCTTGTGCTCTTTCTAGATCTTGT	4862
245	Qy	-----	245
4861	Db	CACGTGAGATCTCCTAGAAAGAGAGAGACATTTCTCCGCTGCTGGAAGAAATGTCAGGAG	4802
245	Qy	-----	245
4801	Db	TTTTGCTAAGTGAATTTAATAAGACTAGTAGTGTCTCAGTAATAATCATCAGTAGGAGA	4742
245	Qy	-----	245
4741	Db	AATCTATATTTTAAAGCTATTTGTAATAGTCACATAATNGCACCCCTTTAAAAAGTCTGAAG	4682
245	Qy	-----	245
4681	Db	TATGGCATTTCAAATTTGTAGCAAACTGTTATCTTTAAGTCATGAGACACTTAAACTCATA	4622
245	Qy	-----	245
4621	Db	CATGAAGTTAACTGCTCTTCTCTTTCTAAGCTGATAAAAGTTACTAAAGACCTTTTAACTA	4562
245	Qy	-----	245
4561	Db	CATCACTGGAAATAATCTTAAATTCGTTAATTCCTCAGAACTCTTTACAGAGAAATGTAT	4502
245	Qy	-----	245
4501	Db	ACCGTGAATGTTGGCAGCTCATTTGGCTTCCTCTCCTTCCCTACTTCTCTAGCATGACCTT	4442
245	Qy	-----	245
4441	Db	GCTCCTTTGTAGAACHCTTTTACTAATGAGCTACAGAAAGATAGGATCCTGGTTGCCA	4382
245	Qy	-----	245
4381	Db	GTCCTTGAGCTAAGTCTAAGTGATTTTGCACAAATGCTCTATTGTTTGTAAATTTCTCTAAC	4322
245	Qy	-----	245
4321	Db	TTTGTGATTAATTTATTTCTTCCTTCTTCATATGTAATGCAAGGTTGATTTGATTTTC	4262
245	Qy	-----	245
4261	Db	ATTACTTGAGCCTGGGGCTCTAGCTACACAGAGCTCTGTGAGCACACTTGGGCTAACTGG	4202
245	Qy	-----	245
4201	Db	TAAGTCTCGAGTCACACTTGC AAAATGAAC TTAAC TTTAAACTGAAGGAGGAAGTCTGCA	4142
246	Qy	-----SerPh	247
4141	Db	TTAGTTGCAATGTTGCTTAGTTGAAATTTGTGACTTCGTTTTTCTCCGTCTTTCTTGCTAGTTT	4082
247	Qy	eAspMetCysProAsnAsnCysSerGlyArgGlyGluCysIysIleSerAsnSerSerGI	267
4081	Db	TGACATGTGTCGGAATAATTTGCTCAGGCGGAGGAGTGTAAAGACGAGTAACACGACGAG	4022

Qy	267	uThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCy	287
Db	4021	CGCTGTTGAGTGTGAATGTTCTGAAACCTGGAAGGGGAGTCTGTGTGACATTTCTCTCACTG	3962
Qy	287	sThrAspAsnCysIlePheProHisArgGlyIleCysAsnSerSerAspValargGlyCy	307
Db	3961	TACAGCAACTGTGGCTTTCTCACCAGGCACTGTAAATGCCAAGCATACCCAGAGGGTG	3902
Qy	307	sSerCysPheSerAspTrpGln	314
Db	3901	CTCTGCTTTCTTCTCACTGGCA-TGGTAGGAGTGGCCCTTCAAAGTCCTTTCCCTCACAGA	3843
Qy	314	-----	314
Db	3842	AAGTTCGAGTTAGTGAACTGTAGGGAGTTATTTTAGAGCCATAGCCTTATGTTAAT	3783
Qy	314	-----	314
Db	3782	CATGTATCAGGTGGCATTAAAGCATTCACTCCATTATTTCAGATTCTTTGAATTTACAT	3723
Qy	314	-----	314
Db	3722	ATATGTGGTTGTATCTATGTGTTAGAAATGTGGAAGGTTTGACAAATCCCTTTGCAAC	3663
Qy	314	-----	314
Db	3662	TCATTCTGGGAATTAGGGAACATGGTCTGGTGATAGAGAAGCAGAGACTCAAGCTGAGA	3603
Qy	314	-----	314
Db	3602	GGAGGACAAGCGTTCTTTAGAGTCTCAGTCTGTCTCATCTAGTGGCATACCCCAACAAC	3543
Qy	314	-----	314
Db	3542	CGGGTACTAAGTACTGTAGGGAATTCCTCGGGGTATTGAGGAAGCAGTCACTCGTTGGTT	3483
Qy	314	-----	314
Db	3482	TAGTGAATTTTCTTCTATGCTTAGTGCTCAGAAAGTAAAGTGTTCATGTAATGTTTTTA	3423
Qy	314	-----	314
Db	3422	GTTTGTCCAATCTCCCTTTGTCTTTTCATCTATTATTCAGCGCTGGACTTGCACAGAAGAT	3363
Qy	314	-----	314
Db	3362	CGATTTTGATCTGTCAATTATTTCAACCCTAATGCCTTTCCACATGAGTTAATAAT	3303
Qy	314	-----	314
Db	3302	AACCCCTTGGTCTACCTTTCTACTAAATGTGTCTTCAGCAGAGATTTTTCCTCTAAAAA	3243
Qy	314	-----	314
Db	3242	AGTTTAACCTGATTATCAGAAATGGCCAAATATACAGAAAGTGTGTAGTATTAAATTTTA	3183
Qy	314	-----	314
Db	3182	TAAGGAGTGAATTGGTTTGGTCTGTATCTTGAGTAAAGTTAAAAACCTCATGTTAGGTGCT	3123
Qy	314	-----	314
Db	3122	AAGCAGATGCCCTTCTGGTTTAAGAGTACTGGTTGGCTTCCAGAGGACTCAGGTTTGATC	3063
Qy	314	-----	314
Db	3062	CTAGCATCCACAAGTAGCTCATAGATCTGTAACTTAGTTCAGGGGATTCATGCGCT	3003
Qy	314	-----	314
Db	3002	TTTTCTGCATCATTTGGGTACCAGGAAGCAAGTGGTATATAGGCATACATGTAAACAAA	2943





```
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 919 GluArg 920
Db 6 GAAAGG 1

RESULT 13
US-09-614-124B-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Tongtong
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; CURRENT FILING DATE: 2001-07-11
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Alignment Scores:
Pred. No.: 8,93e-32 Length: 246
Score: 446.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-614-124B-966 (1-246)
QY 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
Db 246 CTGGGAATATGACGTATCTCAGAGCATGTCACCAAGCTCACCTTAACCCCATGGTCCGC 187
QY 859 LeuArgLysIleAsnValSerTyTrpCysTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTTCCGAAGATCAATGTCTCTACTGCTGGGAGAGATATGTCCTCAATTTACAAATAGT 127
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACAGTGGATCCGCTGAGCCAGTATGCTGGATTCGTGGNAATTTATCAGAA 67
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 919 GluArg 920
Db 6 GAAAGG 1

RESULT 14
US-09-671-325-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
```

```
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Ligu
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; CURRENT FILING DATE: 2000-09-26
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-671-325-966

Alignment Scores:
Pred. No.: 8,93e-32 Length: 246
Score: 446.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.63% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-671-325-966 (1-246)
QY 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
Db 246 CTGGGAATATGACGTATCTCAGAGCATGTCACCAAGCTCACCTTAACCCCATGGTCCGC 187
QY 859 LeuArgLysIleAsnValSerTyTrpCysTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTTCCGAAGATCAATGTCTCTACTGCTGGGAGAGATATGTCCTCAATTTACAAATAGT 127
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACAGTGGATCCGCTGAGCCAGTATGCTGGATTCGTGGAAATTTATCAGAA 67
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 919 GluArg 920
Db 6 GAAAGG 1

RESULT 15
US-08-083-590A-21
; Sequence 21, Application US/08083590A
; Patent No. 5786158
; GENERAL INFORMATION:
; APPLICANT: Artavanis-Tsakonas, S. et al.
; TITLE OF INVENTION: Therapeutic And Diagnostic Methods
; TITLE OF INVENTION: And Compositions Based On No. 5786158ch Proteins And
; TITLE OF INVENTION: Nucleic Acids
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/083,590A
```



D	b		1284	TTGTGAGCATGCAGAAAATGTGTGAACACGGAT---GGCGCCCTTCACACTGTGAGTGTCT	1349
Q	y		598	pATrTpSerValLeuProArgProAspLeuHisAspValasnAargPheGlyHisse	618
D	b		1341	GAAGGGTTATGCAGGACCTCGTTGTGAGATG-----GACATAATCAGTGCT---	1391
Q	y		618	xAlaValLeuHisAsnSerThrMetTyrrVal-----PheGlyClyPheAsnSerLeuLe	636
D	b		1392	AGACCCTGCCAGAATGATGCTACCTGTCTGGTAAGATTGGAGGCTTCACATGTCTGTG	1451
Q	y		636	uLeuSerAspileLeuValPheThrSerGlulnlnCysAspAlahisAargsSerGluala--	655
D	b		1452	CATCCACGGT-----TTCAAAGGTGTCATTTGAAATTAGAAATAAATGAATGCA	1502
Q	y		656	-----AlaCysleualaalaGlyProGlylAargCysvalTrpAsnThrGlyserse	673
D	b		1503	GAGCAACCCCTGTGTGCACAATGGG-----CACTGTGTGATAAAGTCAATCGTTTT	1553
Q	y		673	xGlnCysileSerTrpAlaLeuAlaThrAspGlulnlnGlululysLeulysSerGlucy	693
D	b		1554	CCAGTGCCTGTGTCTCTCTGCTTTCACCTGGGCCAGTTTGCAGATTGCATATTGATGACTG	1613
Q	y		693	sPheSerIysArgThrLeuAspHisAspAargCysAspGlnHisThrAspCysTyrrSerCy	713
D	b		1614	TTCAGACTCCGGTGTCTGAATGGCGCAAGTGTATCGATCACCCGAATGGCTATGAATG	1673
Q	y		713	sThraAlaAsnThrAsnAsp-----CysHistripCysAsnAspHisCysValPr	729
D	b		1674	CCAAGTGGCCACAGTTTCACTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1733
Q	y		729	o-----ArgAsnHisSerCysSerGlugly-----	737
D	b		1734	CGATCCTTGCACCATGTCAGTGTCTCAGGATGGTATTGATTCCTACACCTGCATCTGCAA	1793
Q	y		738	-----GlnIleSerIlePheArgTyrGlusnCyl	747
D	b		1794	TCCCGGTATACATGGCGCCATCTCCAGTGACACAGATTGATGAATGTGTACAGACGCCCTTG	1853
Q	y		747	sProLysAspAsnPromet-----TyrTyrcysAsnlyslysThrse	761
D	b		1854	CCTGAACAGTGTCTGCTGCATTGACCTGGTGCATGGCTACAGTGCAAC-----	1902
Q	y		761	rCysArgSerCysAlaLeuaspGlnAsnCysGlnTrpGluproArgAsnGlnlucysIl	781
D	b		1903	-TGCCACCGGCGCACGTCAGGGGTAAATGT-----GAAATTAATTTTGTATGACTGTGC	1955
Q	y		781	eAlaLeupro-----GlusnilleCysGlylleGlyTrpHisleuValGlyAsnSerCy	799
D	b		1956	AAGTAACCTTGTATCATCGAATCTGTAT-----GGATGGCATTAATCGCTACAGTTGTG	2011
Q	y		799	sLeuLytlethrThrAlalysGlusnTyrrAspAsnAlalysLeuPheCysAargAsnHi	819
D	b		2012	T-----CTGCTCACCGAGATTCA	2029
Q	y		819	sAsnAla-----LeuLeuAlaserLeuthrThr-----	828
D	b		2030	CAGGCGCAGATGPAACATTCATCATGATGATGTGCTTCAATCCCTGTGCGCAAGGGTG	2089
Q	y		829	-----GlnLysLysValGluPheValLeuLysGlnLeuArg-----	840
D	b		2090	CAACATGTPATCAACGGTGAATGGTTTCGCTGTATATGCCCGGAGGACCCCATCACC	2149
Q	y		841	----lleMetGlnSerSerGlnSerMetSerLysLeuThrThrProTrpValGlyLe	859
D	b		2150	CCAGCTGCTACTCACAGGTGAACAAATGCTCGAGCAATCCCTGCATCCATGG-----	2201
Q	y		859	uArgLyslleAsnValSerTyrrTrpCysTrpGluAspMetSerPropheThrAsnSerLe	879
D	b		2202	---AAACTG-----TACTGG-----AG	2215
Q	y		879	uLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlylle-----	895
D	b		2216	GTCTCAGTGAATAA-AGTGTCTC-TGTGATCGAGGTGGTGGTATCGCACTGTGAAGT	2273

Qy	896	-----LeuSerGluProSerThrArgGly-----	903
Db	2274	GGCAAAAAATGAATGCCTTTTCGAATCCATGCCAAGATGGAGAACTGTGACAAATCGGT	2333
Qy	903	-----	903
Db	2334	GAATGATACAGGTGTACTTGCAGAAGAGGCTTTAAAGGCTATAAATGCCAGGTGAATAT	2393
Qy	904	-----LeuLysAlaIaThrCysIleAsnProLeuAsnGI	915
Db	2394	TGATGAATGTCCTCAAAATCATGCCCTGGAACCAAGGAACCTGCTTTTCATGACATAAGTGG	2453
Qy	915	ySerValCysGluArgProAlaAsnHisSerAlaLysGlnCysArgThr-----	931
Db	2454	CTACACTTGCACCTGTGTGTCCTACACAGGCAGAAATTTGTACAGACAGTATTGGTGCC	2513
Qy	932	-----ProCysAlaLeuArgThrAlaCysGlyAsp-----	941
Db	2514	CTGTTCCCAAAACCCCTTGTGAGAAATGCTGTGTTTGCAAAGAGTCACAAATTTTGAGAG	2573
Qy	942	-----CysThrSerGlySerSerGI	948
Db	2574	TTATACTTGTGTGTCTCTGCTGGCAAGGTACAGCGGTACCAATTGCATTTGACGA	2633
Qy	948	uCysMetTrp-----CysSerAsnMetLysGlnCysValAspSer---AsnIaIaTrpVa	965
Db	2634	GTGTATCTCCAAGCCCTGCATGAACCATGGTCTCTGCCATAACACCCAGGCAGACTACAT	2693
Qy	965	IaIaSerPhePro-----	969
Db	2694	GTGTGAATGTCCACACGGCTTCACTGGGATAAGTCCACAGACATGAATGAGTGTCTGAGTGA	2753
Qy	970	-----PheGlyGlnCysMetGluTrpTyrThrMetSerThrCys--	982
Db	2754	TGCCAATCCTTCCACAGAATGGAGTCTCTGTATGGATGGAGTGAATACTTCTCTGCTCC	2813
Qy	982	-----	982
Db	2814	CTGCCTTCCGGGTTTCACTGGGGATAAGTCCACAGACATGAATGAGTGTCTGAGTGA	2873
Qy	983	-ProGluAsn-----CysSerGlyTyR-----CysThrCysSe	994
Db	2874	ACCTGTGAAGAATGGAGGACCTGCTCTGACTAGCTCAACAGTTTACACTTGCAGAGTGCCA	2933
Qy	994	r-----HisCysLeuGluGlnProGlyCysGlyTrpCysThrAspPr	1008
Db	2934	GGCAGGATTTGATGGACTCCATTGT-----GAGAACACATCAATCAGTGCACGTAGAG	2987
Qy	1008	oSerAsnThrThrGlyLysGlyCysIleGluGlySerTyRlysGlyProValysMetPr	1028
Db	2988	CTCCTGTTTCAATGGTGGCAGATGTGTGTATGGG-----	3021
Qy	1028	oSerGlnAlaProThrGlyAsnPheTyRProGlnProLeuAsnSerSerMetCysLe	1048
Db	3022	-----ATTAACTTCTTCTTGCTTT-----	3041
Qy	1048	uGluAspSerArgTyRAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHi	1068
Db	3042	GTGGCCTGTGGGTTTCACTGGATCCTTCTGCCTCCATGAGATCAATGAATGCAGCTCTCA	3101
Qy	1068	sSerLysCysIleAsnGlnSerIleCysGlu-----LysCy	1080
Db	3102	T-----CATGCTGAATGAGGGAACGTGTGTGTATGGCTGGTACCTACCGCTGCAGCTG	3158
Qy	1080	sGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyRGIyAspPr	1100
Db	3159	CCCCCTGGGCTACACTGGGAAAAACTGTCAACCCCTGGTGAATCTTCGAGTCGGTCTCC	3218
Qy	1100	o---ThrAsnGlyClyLysCysGlnProCysLysCysAsnGlyHisIaSerLeuCysAs	1119
Db	3219	ATGTAAAAACAAAGGTACTTGTGTTCAGAAAAAGACAGATGCCAGTGC---CTATGCTCC	3275

1119	nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeu	1139
QY	:::::	
Db	ATCTGGATGGCGTGGCGCTATTGTGCAGTGCCCAATGTCCT	3318
DB	-----	
1139	uCysGluValAluAsnArgfclnglGlyAsnProLeuArgGlyThrCysTyrtYrThr	1158
QY	:::::	
Db	-TTGTACATAGCAGCGCTCCAGACGAGAGTGCTGTTGTGAACACTGTGCCAGCATCA	3375
db	-----	

Search completed: March 1, 2004, 07:12:55  
Job time : 441 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 1, 2004, 05:13:11 ; Search time 11558 Seconds  
(without alignments)  
5358.816 Million cell updates/sec

Title: US-09-787-097-12

Perfect score: 1429

Sequence: 1 WYAAAAATEARLRRTATA.....KSGAVRNKQPPAQPGETCI 1429

Scoring table: OLIGO Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21571516995 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6934743

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlp  
-Q/cgn2\_1/USFT0 spool\_p/US09787097/runat\_26022004\_063014\_23748/app\_query.fasta\_1.1607  
-DB=GenEmbl -QFT0=fastap -SUFFIX=oligo.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITG=bits -START=1 -END=1 -MATRIX=oligo -TRANS=human40.cdi -LIST=100  
-DOCALIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL -OUTFMT=ptc  
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09787097@cgn1.1.7794 @runat\_26022004\_063014\_23748 -NCPUs=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -Fgapop=6  
-Fgapext=7 -Ygapop=60 -Ygapext=60 -DELOP=6 -DELEXT=7

Database : GenEmbl.\*  
1: gb.ba.\*  
2: gb.htg.\*  
3: gb.in.\*  
4: gb.om.\*  
5: gb.ov.\*  
6: gb.pat.\*  
7: gb.ph.\*  
8: gb.pl.\*  
9: gb.pr.\*  
10: gb.ro.\*  
11: gb.sts.\*  
12: gb.sy.\*  
13: gb.un.\*  
14: gb.vi.\*  
15: em.ba.\*  
16: em.fun.\*  
17: em.hum.\*  
18: em.in.\*  
19: em.mu.\*  
20: em.om.\*  
21: em.or.\*  
22: em.ov.\*  
23: em.pat.\*  
24: em.ph.\*  
25: em.pl.\*  
26: em.ro.\*  
27: em.sts.\*  
28: em.un.\*

29: em.vi.\*  
30: em.htg.hum.\*  
31: em.htg.inv.\*  
32: em.htg.other.\*  
33: em.htg.mus.\*  
34: em.htg.pln.\*  
35: em.htg.rod.\*  
36: em.htg.man.\*  
37: em.htg.vrt.\*  
38: em.sy.\*  
39: em.htgo.hum.\*  
40: em.htgo.mus.\*  
41: em.htgo.other.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Query Length	DB ID	Description
1	1267	88.7	3819	9	AF106861 Homo sapi
2	1141	79.8	8589	6	AR164816 Sequence
3	979	68.5	4072	6	AR164817 Sequence
4	874	61.2	3597	9	AF034957 Homo sapi
5	564	39.5	2578	9	AK000356 Homo sapi
6	503	35.2	6370	6	AR164815 Sequence
7	496	34.7	2625	6	AR164818 Sequence
8	452	31.6	5632	9	AB011120 Homo sapi
9	252	17.6	4493	4	AF531101 Bos tauri
10	218	15.3	4284	10	AB062913 Mesocric
11	206	14.4	4313	10	AF119821 Mus muscu
12	206	14.4	8739	10	AB038387 Rattus no
13	206	14.4	8827	6	AR164807 Sequence
14	199	13.9	5683	10	AB093245 Mus muscu
15	193	13.5	6083	10	AF116897 Mus muscu
16	181	12.7	543	6	AX778545 Sequence
17	130	9.1	110000	2	EX004762_0 Mus muscu
18	130	9.1	110000	2	EX004762_1 Continuation (2 of
19	130	9.1	110261	10	AL807401 Mouse DNA
20	130	9.1	335799	2	EX005139 Mus muscu
21	104	7.3	315	4	AF301008 Bos tauri
22	94	6.6	2419	6	AR164813 Sequence
23	92	6.4	1012	6	E26747 Structure a
24	92	6.4	4500	10	AB038388 Rattus no
25	82	5.7	246	6	AR273223 Sequence
26	82	5.7	246	6	AR276804 Sequence
27	82	5.7	246	6	AR407079 Sequence
28	82	5.7	246	6	AX368256 Sequence
29	79	5.5	110000	2	Continuation (3 of
30	79	5.5	110000	2	Continuation (4 of
31	77	5.4	3257	9	AF218903 Homo sapi
32	77	5.4	104907	9	AL132773 Human DNA
33	77	5.4	118540	2	AC013324 Homo sapi
34	77	5.4	167357	2	AC055771 Homo sapi
35	77	5.4	206647	9	AF002898 Homo sapi
36	77	5.4	207433	6	AX326809 Sequence
37	76	5.3	1051	6	AR164814 Sequence
38	74	5.2	3589	9	AF218899 Homo sapi
39	74	5.2	155869	2	AC023420 Homo sapi
40	73	5.1	1207	10	AB06968815 Rattus no
41	73	5.1	1728	9	AF218902 Homo sapi
42	73	5.1	90050	6	AR164810 Sequence
43	73	5.1	189283	10	AL833771 Mouse DNA
44	73	5.1	245589	2	AC094932 Rattus no
45	69	4.8	2018	9	AF218896 Homo sapi
46	68	4.8	1576	9	AF218889 Homo sapi
47	68	4.8	51371	9	AL353193 Human DNA
48	68	4.8	63513	2	AC074061 Homo sapi
49	68	4.8	155869	2	AC023420 Homo sapi
50	64	4.5	3112	9	AF218901 Homo sapi



Db 301 TGTGACCGGCGCTGTCTCAAGCGCGTCTGCTCAACCCCTGGCACCGCGCAGTGGCTGTC 360  
QY 121 ProAlaGlyTyrValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140  
Db 361 CCGCGCGGTGGTGGCGGAGCAATGCCAGCACTGCGGGGCGCTTCAGACTAATCGGA 420  
QY 141 SerSerGlyPheValThrAspGlyProGlyValSerTyrLysThrLysThrLysCysThrTyr 160  
Db 421 TCTTCTGGGTTGTGACAGATGGACCTGGANATTATATACAAAGGAAGTGCAGTGG 480  
QY 161 LeuIleGluGlnProLeuArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180  
Db 481 CTCAATTGAAGGACAGCAAAATAGATAATGAGACTTCGTTTCAATCATTTTCTGTCAGAG 540  
QY 181 CysSerTyrAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200  
Db 541 TGTAGTTGGACCAATTTATGTTTATGATGGGACTCAATTTATGACACCGTGTGCT 600  
QY 201 AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal 220  
Db 601 GCATTTAGTGGCCTCATTTGCTCTGAGAGAGATGGCAATGAGACTGTCCCTCAGGTTGT 660  
QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240  
Db 661 GCCACATCAGGTTATGCTTGTGTCATTTTATGATGCTGCTTATTAATTTGACTGGA 720  
QY 241 PheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys 260  
Db 721 TTTAATATTACTTACAGTTTGTATATGTGTCCAAATAAAGTCTCAGGCGGAGAGTGT 780  
QY 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTyrLysGlyGlu 280  
Db 781 AAGATCAGTAATAAGCAGCAAACTGTGTAATGTGAATGTTCTGAAACCTGGAAGGTGAA 840  
QY 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300  
Db 841 GCATGTGACATCTCTCACTGTACAGCAACTGTGGTTTTCCTCATCAGGCACTGCAAT 900  
QY 301 SerSerAspValArgGlyCysSerCysPheSerAspTyrGlnGlyProGlyCysSerVal 320  
Db 901 TCAAGTGATGTACAGAGATGCTCCTGCTTCTCAGACTGGCAGGCTCCTGGATGTTCA 960  
QY 321 ProValProAlaAsnGlnSerPheTyrThrArgGluGluTyrSerAsnLeuLysLeuPro 340  
Db 961 CCTGTACCAGCTAACAGATCATTTTGGACTCGAGAGGAATATTTCTAACTTAAAGCTCCCC 1020  
QY 341 ArgAlaSerHisLysAlaValAlaAsnGlyAsnIleMetTyrValValGlyGlyTyrMet 360  
Db 1021 AGAGCATCTCATAAAGCTGTGGTCAATGGAAACATTAATGTGGGTTGTTGGAGATATG 1080  
QY 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTyrLeu 380  
Db 1081 TTCAACCACTCAGATATATACATGGTTCTAGCGTATGACTTGTCTTAGGGAGTGGCTT 1140  
QY 381 ProLeuAsnArgSerValAsnValValArgTyrGlyHisSerLeuAlaLeuTyr 400  
Db 1141 CCCTAAACCGTTCTGTGAACATGTGGTTGTAGATATGTCATCTTTGGCATTTATAC 1200  
QY 401 LysAspLysIleTyrMetTyrGlyLysIleAspProThrGlyAsnValThrAsnGlu 420  
Db 1201 AAGGATAAAATTTACATGTATGGAGGAAAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
QY 421 LeuArgValPheHisIleHisAsnGluSerTyrValLeuLeuThrProLysAlaLysGlu 440  
Db 1261 TTGAGAGTTTTTTCATTTTATATGAGTCAATGGGTGTGTTGACCCCTAAGGCAAGGAG 1320  
QY 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460  
Db 1321 CAGTATGCACTGTGTGGCACTCTGCACACATTTGTACACTGAAGAATGGCCGAGTGGTC 1380  
QY 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480

Db 1381 ATGCTGGTCACTCTTTGGTCACTGCGCTCTCTATGGATATATAGCAATGTGCAGGAATAT 1440  
QY 481 AspLeuAspLysAsnThrTyrSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500  
Db 1441 GATTTGTATAGAACACATGGAGTATATTACACACCGGGTGCCTTTGTGCAAGGGGT 1500  
QY 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
Db 1501 TACGGCCATAGCAGTGTTTACGACCATAGGACACAGGCGCTATACGTTTCATGGTGGCTAC 1560  
QY 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540  
Db 1561 AAGGCTTTCACTGGCAATAGTACCGCTTGCAGATGATCTCTACCATATATGATGTGAT 1620  
QY 541 ThrGlnMetTyrThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560  
Db 1621 ACCCAGATGTGACCATCTTAAGCACAGCGCAATTTTCGGTTACTTTCACACAGCTGTG 1680  
QY 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
Db 1681 ATAGTGAGTGGAAACCATGCTGTGTTTGGGGGAAACACACACATATGACATCTATGAGC 1740  
QY 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTyr 600  
Db 1741 CATGGCGCAATGCTTCTCTTCAGATTTCTATGCGCTATGACATTTGCTGTGACGGCTG 1800  
QY 601 SerValLeuProArgProAspLeuHisAspValAsnArgPheGlyHisSerAlaVal 620  
Db 1801 TCAGTGTCTCCAGACCTGATCTCCAGCATGATGTCAACAGATTTGGCCATTTACAGCATC 1860  
QY 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIle 640  
Db 1861 TTACACACACACCATGATGTGTGTTGCTGGTTTCAATAGTCTCTCTCCTCAGGAGCATC 1920  
QY 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660  
Db 1921 CTGTTATTCACTCGGAACAGTGTGATGCGCATCGAGTGAAGCGCTTGTITTAGCAGCA 1980  
QY 661 GlyProGlyIleArgCysValTyrAsnThrGlySerSerGlnCysIleSerTyrAlaLeu 680  
Db 1981 GGACTGTGTATTCGGTGTGTGTGGAAACACAGGGTGTCTCAGTGTATCTCCTGGCGCTG 2040  
QY 681 AlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700  
Db 2041 GCACATGATGAACACAGAGAAAGTTAAATCAGAAATGTTTCCAAAAGAACTCTTGAC 2100  
QY 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720  
Db 2101 CATGACAGATGTGACCCAGCACACAGATTTGTATACAGCTGTACAGCCAAACACCAATGACTGC 2160  
QY 721 HisTyrCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
Db 2161 CACTGGTGCATGACCATTTGTCTCCAGGAACACACAGCTGCTCAGAAAGGCCAGATCTCC 2220  
QY 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760  
Db 2221 ATTTTATAGTATGAGAAATTTGCCCAAGGATAACCCCATGTACTACTGTAAACAAGAAC 2280  
QY 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTyrGluProArgAsnGlnCys 780  
Db 2281 AGCTGCAGAGCTGTGCCCTGGACAGAACTGCCAGTGGAGCCCGGAAATCAGAGATGC 2340  
QY 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTyrHisLeuValGlyAsnSerCysLeu 800  
Db 2341 ATTTGCCCTGCCGAAATATCTGTGGCATTTGCTGGCATTTGGTTGGAACTCATGTTT 2400  
QY 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820  
Db 2401 AAAATTACTTACTGCCAGAGAAATTTATGACAAATGCTAAATTTGTTCTGTAGGAACCAAT 2460  
QY 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840  
Db 2461 GCCCTTTGGCTTCTCTTACAAACCCAGAGAGAGTAGAATTTGCTTAAAGCAGCTGCGA 2520



QY 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860  
DB 2521 ATATGAGTCATCTCAGAGCATGTCCAAGCTCACCTTAACCCATGGGTGGCTTCGG 2580  
QY 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880  
DB 2581 AAGATCAATGTGCTACTGCTGCGGAGATATGTCCCATTTACAAATAGTTTACTA 2640  
QY 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
DB 2641 CAGTGGATGCGTCTGAGCCAGTGAAGTGGATTTCTGGAAATTTATCAGAACCCAGT 2700  
QY 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
DB 2701 ACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGG 2760  
QY 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
DB 2761 CCTGCAAAACACAGTGTCAAGCAGTCCCGACACCAATGTCCCTTGAGGACAGCATGTGA 2820  
QY 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960  
DB 2821 GATTGCACACGCGCAGCTCTGAGTGCATGTGTGCGACGACATGAAGCAGTGTGTGGAC 2880  
QY 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980  
DB 2881 TCCAATGCCATGTGGCTCTCTCCCTTTTGGCCAGTGTATGAATGGTATACGATGAGC 2940  
QY 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000  
DB 2941 ACTGCCCCCTGAAATGTTTTCAGGCTACTGTACTGTAGTCTGTATGCTTGTGGACNACCA 3000  
QY 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020  
DB 3001 GGCTGTGGCTGTGTACTGATCCAGCAATACTGGCAAGGGAATGTCATAGAGGGTTCC 3060  
QY 1021 TyrIleGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040  
DB 3061 TATAAGGACAGTGAAGATGCTTCCGAAAGCCCTACAGGAATTTCTATCCAGACCC 3120  
QY 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060  
DB 3121 CTGCTCAATTCAGCATGTGTCTAGAGGACAGCAGATACAACTGTGTCTTTCTTCACTGT 3180  
QY 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080  
DB 3181 CCAGCTTGCCAAATGCCAACGGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAGTGT 3240  
QY 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100  
DB 3241 GAGAACTGACACAGCAAGCACTCGAGACCTGTATATCTGGCTTCTACGCTGATCCC 3300  
QY 1101 ThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120  
DB 3301 ACCAATGGAGGGAATGTACGCCATCAAGTGCATGGCCAGCGCTCTGTGTCAACACC 3360  
QY 1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140  
DB 3361 AACAGGGCAAGTCTTCTGCACACCAAGGGGTCAAGGGGACGAGTGCAGCTATGT 3420  
QY 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160  
DB 3421 GAGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATCTCTTCTT 3480  
QY 1161 IleAsnTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180  
DB 3481 ATTGACTATCAGTTCACTTTAGTCTATCCAGGAAGATGATCGCTATTATACAGCTATC 3540  
QY 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200  
DB 3541 AATTTTGTGCTACTCTCTGACGAACAAACAGGGAATTGGACATGTTTCATCAATGCTCTC 3600

QY 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220  
DB 3601 AAGAATTTCAACCTCAACATCACCTGGCTGCCAGTTCTCAGCTGGAAACCCAGCTGGA 3660  
QY 1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240  
DB 3661 GAAGAGATGCTGTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTCTCTAAT 3720  
QY 1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260  
DB 3721 GAGAAGTTTGATTTTCGCAACCAACCAATATCATCTTCTTGTATTATGTCAGTAATTC 3780  
QY 1261 ThrTrpProIleLysIleGln 1267  
DB 3781 ACCTGGCCCATCAAAATTCAG 3801  
RESULT 2  
LOCUS AR164816 8589 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 14 from patent US 6274339.  
ACCESSION AR164816  
VERSION AR164816.1 GI:16238025  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE  
1 (bases 1 to 8589)  
Moore,K. and Nagle,D.Lynn.  
METHODS and compositions for the diagnosis and treatment of body  
weight disorders, including obesity  
Patent: US 6274339-A 14 14-AUG-2001;  
JOURNAL Location/Qualifiers  
FEATURES  
source  
1..8589  
/organism="unknown"  
/mol\_type="unassigned DNA"  
ORIGIN  
Alignment Scores:  
Pred. No.: 0 Length: 8589  
Score: 1141.00 Matches: 1341  
Percent Similarity: 99.70% Conservative: 0  
Best Local Similarity: 99.70% Mismatches: 2  
Query Match: 79.85% Indels: 4  
DB: 6 Gaps: 0  
US-09-787-097-12 (1-1429) x AR164816 (1-8589)  
QY 87 AlaAlaValSerGlySerAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
DB 256 CGCGCGCGGTGTCCGCTCGCGCGCAGCGCCAGGAATGTGACCGCGCGGTGTGTC 315  
QY 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTyrValGly 126  
DB 316 AACCGCGGTGCTGCAACCTTGGCCAGCGCGCTGTGCTGCCCGCGCGGTGGC 375  
QY 127 GluGlnCysGlnHisCysGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146  
DB 376 GAGCAATGCCAGCACTCGCGGGCGCGCTCAGACTAACTGGATCTTCTGGGTTGTGACA 435  
QY 147 AspGlyProGlyAsnTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166  
DB 436 GATGGACTGGAAATTAATAACAAACGAAGTGCAGTGGCTCATTTGAGGACAGCCA 495  
QY 167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186  
DB 496 AATAGAATAATGACATCTGTTTCAATCAATTTGTACAGAGTGTAGTTGGACCACTTA 555  
QY 187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaPheSerGlyLeuIle 206  
DB 556 TATGTTTATGATGGGACTCAATTTATGCACCGCTAGTGTGCTGCTTATGCTGCTCAT 615  
QY 207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226

616 GTTCTCTGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675  
Db  
227 LeuLeuHisPhePheSerAspAlaLalaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246  
Qy  
676 TTGCTGCATATTTTATAGTCATCTCTTATAATTTGACTGGATTTAAATATTACTTACAGT 735  
Db  
247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266  
Qy  
736 TTTGATATGTCTCCAAATTAACCTGCTCAGCCCGAGGAGTGTAAGATCAGTAATAGCAGC 795  
Db  
267 Glu-ThrValGluCysGluCysSerGluAsnTrpIysGlyGluAlaCysAspIleProHi 286  
Qy  
796 GA-TACTGTGTAATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCA 854  
Db  
286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgG1 306  
Qy  
855 CTGTACAGCAACCTGTGGTTTCTCATCGAGGATCTGCAATCAAGTGATGTGAGAGG 914  
Db  
306 YCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnG1 326  
Qy  
915 ATGCTCTCTCTCAGACTGGCAGGCTCTGGATGTTTCAGTTCCTGTACCACTTAACCA 974  
Db  
326 nSerPheTrpThrArgGluGluTyrSerAsnLeuIysLeuProArgAlaSerHisLysAl 346  
Qy  
975 GTCAATTTGGACTCGAGAGGAATATCTAACTTAAAGCTCCCGAGCATCTCATAAAGC 1034  
Db  
346 aValValAsnGlyAsnIleMetTrpValValGlyTyrMetPheAsnHisSerAspTy 366  
Qy  
1035 TGTGGTCAATGGAACATTTATGTGGTGTCTGGAGATATATGTTCAACCACTCAGATTA 1094  
Db  
366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386  
Qy  
1095 TAAACATGGTCTAGCGTATGAGCTTGTCTTAGGAGTGGCTTCCACTAAACCGTCTGT 1154  
Db  
386 lAsnAsnValValAlaArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406  
Qy  
1155 GAACAATGGTGTGTAGATATGCTCATCTTTGGCAITATACAGGATAAATTTTACAT 1214  
Db  
406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426  
Qy  
1215 GTATGGAGGAAAAATTTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACA 1273  
Db  
426 lHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446  
Qy  
1274 TTCAATATGAGTCATGGGTGTGTGACCCCTTAAGCAAGAGCAGATGTCAGTGGTGTG 1333  
Db  
446 lYHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466  
Qy  
1334 GGCACCTCTGCACACATTTGACACTGAAGATGGCCGAGTGTCTATGCTGGTCACTTTG 1393  
Db  
466 lYHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486  
Qy  
1394 GTCACCTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAACA 1453  
Db  
486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSerSerV 506  
Qy  
1454 CATGGAGTATATTACACACCCAGGTCCTTGTGCAGGGGTTACGGCCCATACAGTGTG 1513  
Db  
506 aTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526  
Qy  
1514 TTATACACCATAGGACAGGGCCCTATAGCTTCATGGTGGCTACAAGGCTTTTCAGTGCCA 1573  
Db  
526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546  
Qy  
1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGATACCCAGATGTGAGCCA 1633  
Db  
546 lLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrm 566  
Qy  
1634 TTCTTAAGGACACCCGATTTTCCGTTACTTGCACACAGCTGTGATAGTGTGAGTGAACA 1693  
Db  
566 etLeuValPheGlyG1YAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586  
Qy  
1694 TGCTGTGTGTTGGGGGAAAAACACACAAATGACACATCTATAGCCATGCGCCCAATGCT 1753  
Db

586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTTrpSerValLeuProArgP 606  
Qy  
1754 TCTCTTCAGATTTTCATGGCTATGACATTCCTGTGTGACCTGGTCACTGCTTCCCAGAC 1813  
Db  
606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrm 626  
Qy  
1814 CTGATCTCCACCATGATGTCAACAGATTTGSCCATTCAGCAGTCTTACACACAGCACCA 1873  
Db  
626 etTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerG 646  
Qy  
1874 TGTATGTGTTCGGTGGTTTCAATAGTCTCTCTCTCAGCAGCATCTCTGGTATTACCTCGG 1933  
Db  
646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
Qy  
1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTAGCAGCAGGACCTGGTATTCCGT 1993  
Db  
666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686  
Qy  
1994 GTGTGTGAAACACAGGTCGTCTCAGTGTATCTCTGTGGGCGCTGGCAACTGATGAACAAG 2053  
Db  
686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706  
Qy  
2054 AAGAAAAGTTAAATACAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTACC 2113  
Db  
706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726  
Qy  
2114 AGCACACAGATTTGACAGCTGCACAGCAACCAATGACTGCCACTGGTGCATGACC 2173  
Db  
726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrcluA 746  
Qy  
2174 ATTGTGTCCCGAGAAACACAGCTGCTCAGAAGCCAGATCTCCATTTTATGGTATGAGA 2233  
Db  
746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766  
Qy  
2234 ATTGCCCCAAGGATAACCTATGTACTACTGTAAACAAGAGAGACCAGCTGCAGGAGCTGTG 2293  
Db  
766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786  
Qy  
2294 CCCTGGACCAAACTGCGAGTGGGAGCCCGGAATCAGGAGTGCATTCCTCCCTGCCGAAA 2353  
Db  
786 snIleCysGlyIleGlyTyrHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaI 806  
Qy  
2354 ATATCTGGCATTTGGCTGGCATTTGGTTGGAACACTCATGTTTGAATAATTACTACTGCCA 2413  
Db  
806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826  
Qy  
2414 AGGAGAATTATGACATGTCTAAATTTGTTGTAGGAACCAACAATGCCCTTTTGGCTTCTC 2473  
Db  
826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846  
Qy  
2474 TTACAACCCAGAAAGAGTGAATTTGCTCTTAAGCAGCTGCGAATAATGAGTCATCTC 2533  
Db  
846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866  
Qy  
2534 AGACCATGTCAAGCTCACCTTAACCCCATGGGTGGCCCTTCGAAGATCAATGTGTCT 2593  
Db  
866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886  
Qy  
2594 ACTGGTCTGGGAAGATATGTCCCAATTACAAATAGTTTACTACAGTGGATGCCGCTG 2653  
Db  
886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906  
Qy  
2654 AGCCCACTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTTCGGGACTGAAGG 2713  
Db  
906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926  
Qy  
2714 CTGCAACCTGTCATCAACCCACTCAATGGTGTCTGTGAAAGGCCCTGCACCAACACAGTG 2773  
Db  
926 laLysGlnCysArgTrpProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946  
Qy  
2774 CTAAGCAGTCCCGACACCACTGTCCTTTAGGACAGCATGTGGAGATTGCACACAGCGCA 2833  
Db

QY 946 erSerGluCysMetTyrCysSerAsnMetIleGlnCysValAspSerAsnAlaTyrVala 966  
 DB 2834 GCTCTCAGTGCATGTGGTCAGCAACATGAAGCAGTGTGGACTCCCAATGGCTATGTGG 2893  
 QY 966 laSerPheProPheGlyGlnCysMetGluTyrThrMetSerThrCysPProGluA 986  
 DB 2894 CCTCCTTCCCTTTTGGCCAGTGATGGATGGTATACGATGAGCAGCTGCCCTGAA 2953  
 QY 986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTyrCysT 1006  
 DB 2954 ATTGTTTCCAGCTACTGCTAGCTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGTA 3013  
 QY 1006 hrAspProSerAsnThrGlyLysCysIleGluGlySerTyrLysGlyProVal 1026  
 DB 3014 CTGATCCAGCAATATCTGGCAAGGGAATGATAGAGGTTCTTATAAGGACAGTGA 3073  
 QY 1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046  
 DB 3074 AGATGCTTTCGCAAGCCCTTACAGGAATTTCTATCCAGAGCCCTGCTCAATTCAGCA 3133  
 QY 1046 etCysLeuGluAspSerArgTyrAsnThrSerPheIleHisCysProAlaCysGlnCysA 1066  
 DB 3134 TGTGCTTAGAGGACAGCAGATCACTGGTCTTTCATTCCTGCTCCAGCTTCCCAATGCA 3193  
 QY 1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086  
 DB 3194 ACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCCTGACCACAG 3253  
 QY 1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyLysC 1106  
 DB 3254 GCAAGCACTCGAGAGCTGCATCTCGGCTTCTACGGTATCCCAATGAGGGAAT 3313  
 QY 1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126  
 DB 3314 GTAGCCATGCAAGTGCATGGGACGCGTCTGTGCAACCAACACACGGGCAAGTGTCT 3373  
 QY 1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146  
 DB 3374 TCTGCACCAACAGGGCGTCAAGGGGAGGAGTCCAGCTATGTGAGGTAGAAAATCGAT 3433  
 QY 1146 yrGlnGlyAsnProLeuArgGlyThrCysTyrThrLysLeuLeuIleAspTyrGlnPheT 1166  
 DB 3434 ACCAAGAAACCTCTCAGAGGAACATGTTATATATCTCTTCTATGACTATCAGTTCA 3493  
 QY 1166 hrPheSerLeuSerGlnGluAspAspArgTyrThrLysAlaIleAsnPheValAlaThrP 1186  
 DB 3494 CCTTTAGTCTATCCAGGAAGATGATCGCTATTTACACAGCTATCAATTTTGGCTACTC 3553  
 QY 1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206  
 DB 3554 CTGACCAACCAACAGGATTTGGACATGTTCATCAATGCTCCCAAGAAATTTCAACCTCA 3613  
 QY 1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal 1226  
 DB 3614 ACATCAGCTGGCTGCGAGTTTCTCAGCTGGAACCCAGGCTGAGAGAAGATGCTGTGTG 3673  
 QY 1226 alSerLysThrAsnIleLysGluTyrIleAspSerPheSerAsnGluLysPheAspPheA 1246  
 DB 3674 TTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTATGAGAAGTTTGAATTTTC 3733  
 QY 1246 rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIleLysI 1266  
 DB 3734 GCACCAACCAACATATCATTCTTTGTTATGTCAGTAATTTTCACTTGGCCCATCAAA 3793  
 QY 1266 leGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrP 1286  
 DB 3794 TTCAGATGCTCTCTCAGCAGCAATTTTATGACCTGGTGTACAGTTCTTCTGCTGACTT 3853  
 QY 1286 hePheSerCysPheLeuSerLeuLeuValAlaValValTyrLysIleLysGlnS 1306  
 DB 3854 TCTTCAGTTGTTCTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3913  
 QY 1306 erCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerA 1326

DB 3914 GTTGTGGGCTCCAGAGCTAGAGCAACTTCTTCAGAGATGCAACAGATGGCCAGCC 3973  
 QY 1326 rgProPheAlaSerValAsnValAlaLeuLeuThrAspGluGluProProAspLeuIleG 1346  
 DB 3974 GTCCCTTTGCTCTGTAAATGTCCTTGGAAACAGATGAGGAGCCTCTCTGATCTTATTG 4033  
 QY 1346 lyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLysA 1366  
 DB 4034 GGGGAGGTATAAAGACTTTCCCAACCCATTGCACTGGAGCCGTGTTTGGCAACAAAG 4093  
 QY 1366 laAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyLysLeuProProG 1386  
 DB 4094 CCGCTGCTCTCTGTGTTGTGAGGCTCCTCGAGGCTGGGTGGCATCCCTCTCTCTG 4153  
 QY 1386 lyGlnSerGlyLeuAlaValAlaSerAlaLeuValAlaPheIleSerGlnGlnMetProIleV 1406  
 DB 4154 GGCAGTCAGGCTTGTGTCGCCAGCCCTGTGGGACATTTCTCAGCATGCCGATAG 4213  
 QY 1406 alTyrLysGluLysSerGlyValaValArgAsnArgLysGlnGlnProProAlaGlnProG 1426  
 DB 4214 TGTACAGGAGAGTCAAGAGCGGTGAGAAACCGAGAGCAGAGCCCTGTCACAGCCTG 4273  
 QY 1426 lyThrCysIle 1429  
 DB 4274 GGACCTGCATC 4284  
 RESULT 3  
 ARI164817  
 LOCUS ARI164817 4072 bp DNA linear PAT 17-OCT-2001  
 DEFINITION Sequence 16 from patent US 6274339.  
 ACCESSION ARI164817  
 VERSION ARI164817.1 GI:16238027  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 4072)  
 AUTHORS Moore,K. and Nagle,D.Lynn.  
 TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity  
 JOURNAL Patent: US 6274339-A 16 14-AUG-2001;  
 FEATURES Location/Qualifiers  
 source 1..4072  
 /organism="unknown"  
 /mol\_type="unassigned DNA"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 0 Length: 4072  
 Score: 979.00 Matches: 1179  
 Percent Similarity: 99.66% Conservative: 0  
 Best Local Similarity: 99.66% Mismatches: 2  
 Query Match: 68.51% Indels: 4  
 DB: 6 Gaps: 0  
 US-09-787-097-12 (1-1429) x ARI164817 (1-4072)  
 QY 87 AlalaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
 DB 256 CGCGCGGCTGCTCGGCTCGCGCGCAGCGAGCCAAAGAAATGTGACCGCCGCTGTGTC 315  
 QY 107 AsnGlyGlyArgCysAsnProGlyThrGlyLncysValCysProAlaGlyTTPValGly 126  
 DB 316 AACGGCGTGCCTGCAACCTTGGACCGCCAGTGGCTCTGCCCGCCGCTGGTGGGC 375  
 QY 127 GluGlnCysGlnHisCysGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146  
 DB 376 GAGCAATGCCAGCACTCGGGGGCGCTTCAGACTAACTGGATCTTCTGCGTTTGTGACA 435  
 QY 147 AspGlyProGlyAsnTyrLysTyrLysCysThrTrpLeuIleGluGlyGlnPro 166  
 DB 436 GATGAGCTGGAATTAATAAACAAGAGTGACGCTGGCTCATTTGAGGACAGCCA 495

QY 167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisIleu 186  
Db 496 AATAGAAATATGAGACTTCGTTTCAATCAATTTTGTACAGAGTAGTGGACCATTTA 555  
QY 187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206  
Db 556 TATGTTTATGATGGGAGACTCAATTTATGACCCGCTAGTTGCTGCAATTTAGTGGCTCAT 615  
QY 207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226  
Db 616 GTTCTCGAGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675  
QY 227 LeuLeuHisPhePheSerAspAlaIleTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246  
Db 676 TTGCTGCATTTTATGATGCTGCTTATATTTGACTGGATTTAAATATTACTTACAGT 735  
QY 247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266  
Db 736 TTTGATATGTGTCCTCAATTAATGCTCAGCCGAGGAGAGTGAAGATCAGTAATAGCAGC 795  
QY 267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286  
Db 796 GA-TACTGTTGAATGGAATGTTCTGAAAACCTGGAAAGGTGAAGCAATGACATTCCTCA 854  
QY 286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgI 306  
Db 855 CTGTACAGACAACTGTGGTGTTCCTCATCGAGGCATCTGCAATTCAGAGTATGTCAGAGG 914  
QY 306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnG 326  
Db 915 ATGCTCCTGCTTCTCAGACTGGCAGGCTCTGGATGTTTCAGTTCCTGTACAGCTAAGCA 974  
QY 326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346  
Db 975 GTCAATTTGCACTCGAGAGGAATATTTCTAACTTAAAGCTCCCGACAGCATCTCATAAAGC 1034  
QY 346 aValValAsnGlyAsnIleMetTrpValValGlyTyrMetPheAsnHisSerAspTyr 366  
Db 1035 TGTGGTCAATGGAACATTTATGTGGTGTCTGGAGGATATATGTTCAACCACTCAGATTA 1094  
QY 366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTyrLeuProLeuAspArgSerVa 386  
Db 1095 TAAACATGGTCTTAGCGATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCGTTCGT 1154  
QY 386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406  
Db 1155 GAACAATGTGGTGTGTAGATATGGTCATTTTGGCATTTATACAAGGATATAAATTTACAT 1214  
QY 406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426  
Db 1215 GTATGAGGAGAAATTTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACA 1273  
QY 426 lHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446  
Db 1274 TTCATANTGATCATGGGTGTGTGACCCCTAAGCCAAAGAGCAGATGACAGTGGTGTG 1333  
QY 446 lYHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466  
Db 1334 GGCACCTCTGCACACATTTTACATGGAAGATGGCCGAGTGTGTATGCTGGTCACTTTTG 1393  
QY 466 lYHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486  
Db 1394 GTCACTGCCCTCTATGATGATATATAGCAATGTGCGAGGAATATGATTTGGATAGAACA 1453  
QY 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerV 506  
Db 1454 CATGGAGTATTTACACACCCAGGGTGGCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513  
QY 506 aITyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526  
Db 1514 TTTTACGACCATAGGACCGGGCCCTATACGTTTCATGCTGGCTACAGGCTTTTCAGTGCCA 1573

QY 526 snLysTyrArgLeuAlaAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546  
Db 1574 ATAAGTACCGGCTTGCAGATGATCTCTACGATATGATGTGGATACCAAGATGGACCA 1633  
QY 546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566  
Db 1634 TTCTTAAAGGACAGCCGATTTTTCGTTTACTTGCACACAGCTGTGATAGTGGTGAACCA 1693  
QY 566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586  
Db 1694 TGCTGGTGTTTGGGGGAAACACACACAAATGACACATCTATGAGCCATGGGCCCAATGTCT 1753  
QY 586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606  
Db 1754 TCTCTTCAGATTTTCATGCGCTATGACATTCCTGTGACCGCTGTGCTGCTTCCAGAC 1813  
QY 606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626  
Db 1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTTCAGCAGTCTTTACACAAACAGCACA 1873  
QY 626 etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646  
Db 1874 TGTATGTGTTTCGGTGGTTCATAGTCTCTCTCTCAGCGACATCTCTGGTATTCACCTCG 1933  
QY 646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
Db 1934 AACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTTCCTCAGCAGGACCTGTGTATTCGGT 1993  
QY 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686  
Db 1994 GTGTGTGAACACAGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053  
QY 686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706  
Db 2054 AAGAAAAGTTTAAATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113  
QY 706 lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726  
Db 2114 AGCACAGATTTGTTCAGCTGCACAGCAACCAATGACTGCTCATCTGGTGCATGACC 2173  
QY 726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746  
Db 2174 ATTGTGTCCCAGAAACCAACAGCTGCTCAGAGCCAGATCTCCATTTTAGTATGAGA 2233  
QY 746 snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766  
Db 2234 ATTCCGCCAAGATAACCCCTATGTACTACTGTAAACAAGAACCCAGCTGCAGAGCTGTG 2293  
QY 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786  
Db 2294 CCTGGACCAAGACTGCAGTGGAGGCCCGGAATCAGGAGTGCATTCGCCCTGCCCGAAA 2353  
QY 786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806  
Db 2354 ATATCTGTGGCATTTGGTGGCATTTGGTGGAAACTCATGTTTGAATAATTACTACTGCCA 2413  
QY 806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826  
Db 2414 AGGAGATTATGACAAATGTCTAAATTTGTCTGTAGGAACCAACCAATGCCCTTTTGGCTTCTC 2473  
QY 826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846  
Db 2474 TTACAACCCAGAGAAGTAGAATTTGTCCTTAAGCAGCTGCCGAATAATCAGTCATCTC 2533  
QY 846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866  
Db 2534 AGACATGTCCAAGCTCACCTTAACCCCATGGGTGGCCCTTCGGAAAGATCAATGTGTGCTC 2593  
QY 866 yITpCysTyrGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886  
Db 2594 ACTGGTCTGGAGAGATATGTCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCTG 2653  
QY 886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906

Db	2654	AGCCAGTGTGATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGGACATGAAGG	2713
Qy	906	laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA	926
Db	2714	CTGCACCTGTCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTGCAACCCACAGTG	2773
Qy	926	laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS	946
Db	2774	CTAAGCAGTCCCGACACCATGTGCTTGTGAGACAGCATGTGAGATTTCACACCGCGCA	2833
Qy	946	erSerGluCysMetTyrCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA	966
Db	2834	GCTCTCAGTGCATGTGGTCAGCAACATGAGCAGTGTGTGACTCCAATGCTATGTGG	2893
Qy	966	laSerPheProPheGlyGlnCysMetGluTyrPyrThrMetSerThrCysProProGluA	986
Db	2894	CCTCCTTCCCTTTTGGCCAGTGTATGCAATGGTATACGATGAGCACCTGCCCCCTGAA	2953
Qy	986	snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTyrCysT	1006
Db	2954	ATTGTTTCAGCTACTGTACTGTAGTCATGTCTGAGCAACAGAGCTGTGCTGTGTGA	3013
Qy	1006	hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL	1026
Db	3014	CTGATCCAGCATACTGGCAAGGGAATGCATAGAGGTTCTCTATATAAGGACCACTGA	3073
Qy	1026	ysKetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM	1046
Db	3074	AGATGCTTCCGAAGCCCTTACAGGAATTTCTATCCAGAGCCCTGTCTCAATTCACGCA	3133
Qy	1046	etCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCysProAlaCysGlnCysA	1066
Db	3134	TGTGCTTAGAGGACAGCATCACTGCTCTTCACTTCATCTGTCAGCTTGCCATGCA	3193
Qy	1066	snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG	1086
Db	3194	ACGGCCACAGTAAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCCTGACCA	3253
Qy	1086	lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC	1106
Db	3254	GCAAGCACTCGAGAGCTGCATATCTGCTTCTACGGTATCCCAATGAGGGAAT	3313
Qy	1106	ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP	1126
Db	3314	GTGAGCATGCAAGTCAATGGCCAGCGCTCTCTGTGCAACCAACCAACCGGCAAGTGT	3373
Qy	1126	heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT	1146
Db	3374	TCTGCACCAACAGGGCGTCAAGGGGAGCAGTGCAGCTATGTGAGGTAGAAATTCGAT	3433
Qy	1146	yrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT	1166
Db	3434	ACCAAGAAACCTCTCAGAGGAACATGTTATTATCTCTCTTATTGACTATCAGTTCA	3493
Qy	1166	hrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheValAlaThrP	1186
Db	3494	CTTTTAGTCTATCCAGGAGATGATCGCTATTATACAGCTATCAATTTTGGGTACTC	3553
Qy	1186	roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA	1206
Db	3554	CTGACGAAACAAACAGGATTTGACATGTTCAATCAATGCCITCAAGAAATTTCACTCA	3613
Qy	1206	snIleThrTyrAlaIleSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV	1226
Db	3614	ACATCCTGGCTGCGCAGTTTCTCAGCTGAAACCCAGGCTGAGAAAGATGCTCTGTG	3673
Qy	1226	alSerLysThrAsnIleLysGlyTyrLysAspSerPheSerAsnGluLysPheAspPheA	1246
Db	3674	TTTCAAAACCAACATTAAAGAGTACAAAGATAGTTTCTCTAATGAGAAGTTGATTTC	3733
Qy	1246	rgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTyrProIleLysI	1266

Db	3734	GCAACACCAAAATATACACTTTCTTCTTTATGTACAGTAATTTACCTGGCCCAAAA	3793
Qy	1266	leGln 1267	
Db	3794	TTCAG 3798	
RESULT 4			
AF034957	3597 bp	mRNA	linear PRI 30-SEP-1998
LOCUS	Homo sapiens secreted T cell activation protein Attractin		
DEFINITION	(attractin) mRNA, complete cds.		
ACCESSION	AF034957		
VERSION	AF034957.1		
KEYWORDS	GI:3676346		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1. (bases 1 to 3597)		
AUTHORS	Duke-Cohan, J.S., Gu, J., McLaughlin, D.F., Xu, Y., Freeman, G.J. and Schloessman, S.F.		
TITLE	Attractin (DPPT-1), a member of the CUB family of cell adhesion and guidance proteins, is secreted by activated human T lymphocytes and modulates immune cell interactions		
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 95 (19), 11336-11341 (1998)		
MEDLINE	98409658		
PUBMED	9736737		
REFERENCE	2. (bases 1 to 3597)		
AUTHORS	Duke-Cohan, J.S., Gu, J., Ao, Z., McLaughlin, D.F., Freeman, G.J. and Schloessman, S.F.		
TITLE	Direct Submission		
JOURNAL	Submitted (19-NOV-1997) Division of Cancer, Immunology and AIDS, Dana Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA		
FEATURES	Location/Qualifiers		
source	1..3597		
gene	/organism="Homo sapiens"		
CDS	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	1..3597		
	/gene="attractin"		
	/gene="attractin"		
	/note="secreted T cell activation protein; Method: conceptual translation with partial peptide sequencing; DPPT-1"		
	/codon_start=1		
	/product="Attractin"		
	/protein_id="AAC61902.1"		
	/db_xref="GI:3676347"		
	/translation="MVAAATAEARLRRTAATAALAGSGGPHCVNGECNPGTQC VCPAGWGEQCQHCGRFRTSGSVFTDGNKYKTKCTWLIEGQPNRMLRENH FATECSWDHLVYVDGSIYAPLVAASGLIVPERDNETPEVVAATSGVALLHFFSDA AYNLGTGNIYTFDMPNNGRGECKISNSDTECESENWKGACDIPHTDNGC PFRGICNSSDVRGSCFSDWQPGGCVPPANQSVNRYRGLSLAYKDKIYWG NIMWGVGVNHSDDYNWLYADLASREWLPLNRSVNVVRYRGLSLAYKDKIYWG GKIDSTGNTVELRYFHINESVLLTPKAQYAVVGHSAHLTKNGRVNMLVIFG HCPYGIYSINQYEDLDKNTNTSLHTKQALVQSGYGHSSVYDRTALYVHGKYAF ANKYRLADDYRYVDVTQMTILKSRFRYLTIVSGTLMVFGNTHTDTSMSHG /protein_id="AAC61902.1"		
	/db_xref="GI:3676347"		
	/translation="MVAAATAEARLRRTAATAALAGSGGPHCVNGECNPGTQC VCPAGWGEQCQHCGRFRTSGSVFTDGNKYKTKCTWLIEGQPNRMLRENH FATECSWDHLVYVDGSIYAPLVAASGLIVPERDNETPEVVAATSGVALLHFFSDA AYNLGTGNIYTFDMPNNGRGECKISNSDTECESENWKGACDIPHTDNGC PFRGICNSSDVRGSCFSDWQPGGCVPPANQSVNRYRGLSLAYKDKIYWG NIMWGVGVNHSDDYNWLYADLASREWLPLNRSVNVVRYRGLSLAYKDKIYWG GKIDSTGNTVELRYFHINESVLLTPKAQYAVVGHSAHLTKNGRVNMLVIFG HCPYGIYSINQYEDLDKNTNTSLHTKQALVQSGYGHSSVYDRTALYVHGKYAF ANKYRLADDYRYVDVTQMTILKSRFRYLTIVSGTLMVFGNTHTDTSMSHG AKPFSDFMAYDIACDRWSVLPDSTMTDLAI PAVLHNSMTYVFGFNSLLSDI LVFTSECDADRSEAEACLAAGPGTRCVMTNGSSQICISWALATDEQEKLSKSEF SKRT LDHRCQHTDCYCTANTNDCHNDHCNDRNHSSEGOISIFRYENCPKDNPMYC NKDTCSCALDONCOWEPNOCIALPENI CGIHWLVGNSCLKITAKENDNAKL FCNNHALLASLTQKVEFVLKQIRIMOSSQSMSKLTLPWVGLRKINYSVCWBDM SPFTNSLQWMPSPDAGFCGILSPSTRGLAAATCINPLNGSVCEERPAHNSAKCR TPCALRTACQDCSTGSECCNMWCKQVDSNAYVASFPFGQCNEMTWTCTCPENC GYCTCHLQPGCGCTDPSNCKGCKEYKGPVKPSPQATGNFYDQPLNNSM CLEDSRYNWSFTHCPACQCNHSGKINQISICERCNLTGKHGCTCISGFYGDPTNGG KCOPCKNGHAGLCNTNTGKPCFTTKGVGDEQOLCEVENRYQGNPLRGTCFYATLLID YQFTFSLSQSDRDYTAINFVATPDQNRDLDMFINASKNFNLNIWAASFSAGTOAG EEMPVSVKTIKEYKDSFSEKEDFRNHNITFFVYVSNTFIKIQVQTEQ"		

ORIGIN

## Alignment Scores:

Pred. No.: 0 Length: 3597  
Score: 874.00 Matches: 1160  
Percent Similarity: 99.49% Conservative: 0  
Best Local Similarity: 99.49% Mismatches: 3  
Query Match: 61.16% Indels: 6  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AF034957 (1-3597)

```
QY 105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124
DB 91 TGTGTCAACGGCGGTGCTGCAACCTGCGCACGGCCAGTGCGTCTGCCCGCGCGTGG 150
QY 125 ValGlyGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPhe 144
DB 151 GTGGCGAGCAATGCGAGCACTGCGGGGCCGCTTCAGACTAACTGATCTTCTGGGTTT 210
QY 145 ValThrAspGlyProGlyAsnTyLysTyLysThrLysCysThrTrpLeuIleGluGly 164
DB 211 GTGACAGATGGACCTGGAAATTTATAAATACAAAACGAAGTGCACGTGGCTCAATTGAAGA 270
QY 165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184
DB 271 CAGCCAAATAGAAATATAGACTTCGTTTCAATCATCTTTTGTCTACAGAGTGTAGTGGAC 330
QY 185 HisLeuTyLysValTyAspGlyAspSerIleTyLysAlaProLeuValAlaAlaPheSerGly 204
DB 331 CATTTATATGTTTATGATGGGACTCAATTTATGACCCGCTAGTTGCTGCAATTTAGTGGC 390
QY 205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValAlaAlaThrSerGly 224
DB 391 CTCATTGTTCTTGAGAGAGATGGCAATGAGACTGTCCTCGAGGTTGTTGGCCACATCAGGT 450
QY 225 TyrAlaLeuLeuHisPhePheSerAspAlaAlaTyLysAsnLeuThrGlyPheAsnIleThr 244
DB 451 TATGCTTCTGCTGATTTTATGATGCTGCTTATTAATTTGACTGGATTTAATATTACT 510
QY 245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264
DB 511 TACAGTTTGTATGTGTCCAAATTAATCTCTCAGCCCGAGGAGTGTAAAGTCAAGTAAAT 570
QY 265 SerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIle 284
DB 571 AGCAGCGA-TACTGTTGGAATGTGAATGTTCTGAAAACCTGAAAAGGTGAAGCATGTGACAT 629
QY 284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304
DB 630 TCTCTACTGTACAGCAACTGTGGTTTCTCATCGAGGCATCTGCAATTCGAAGTGATGT 689
QY 304 lArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAl 324
DB 690 CAGAGGATGCTCTGCTTCTCAGACTGGCAGGCTCTGGATGTTCAAGTTCCGTACCAGC 749
QY 324 aAsnGlnSerPheTrpThrArgGluGluTyLysSerAsnLeuLysLeuProArgAlaSerHi 344
DB 750 TAACCACTCATTTGACTCGAGAGAAATTTCTAACTTAAAGCTCCCGAGCATCTCA 809
QY 344 sIysAlaValAlaAsnGlyAsnIleMetTrpValValGlyGlyTyLysMetPheAsnHisSe 364
DB 810 TAAAGCTGTGGTCAATGGAACATTTATGTGGGTGTTGGAGGATATATGTTCAACCACTC 869
QY 364 rAspTyLysMetValLeuAlaTyLysAspLeuAlaSerArgGluTrpLeuProLeuAsnAr 384
DB 870 AGATTATAACATGGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCCACTAAACCG 929
QY 384 gSerValAsnAsnValValValArgTyLysIleSerLeuAlaLeuTyLysAspLysIle 404
DB 930 TTCTGTGAACAATGTGGTTGTTAGATATGTGTCATCTTTGGCATTTATACAGGATAAAT 989
QY 404 eTyLysMetTyLysGlyIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424
DB 990 TTACATGTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGATGATTTGAGATT 1048
```

```
QY 424 heHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlyTyLysVal 444
DB 1049 TTCACATTTCATTAATGACTCATGGTGTGTGTGCCCCCTAAGGCAAGAGCAGTATGCAG 1108
QY 444 alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464
DB 1109 TGGTTGGGCACTCTGCAACATTTGTACACTGAAGAATGGCCGAGTGGTCTATGCTGTCA 1168
QY 464 lePheGlyHisCysProLeuTyLysTrpIleSerAsnValGlnGluTyLysAspLeuAspL 484
DB 1169 TCTTTGGTCACTGCGCTCTCTATGGATATATAAGCAATGTGCAAGAAATATGATTTGATA 1228
QY 484 yAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyLysHis 504
DB 1229 AGAACACATGAGATATATTACACCCAGGGTGGCTTTGTGCAAGGGGGTTAGCGCCATA 1288
QY 504 erSerValTyLysAspHisArgThrArgAlaLeuTyLysValHisGlyGlyTyLysAlaPhe 524
DB 1289 GCAGTGTTTACGACCATAGGACCCAGGCGCTATAGTTCATGGTGGCTACAGGCTTTCA 1348
QY 524 erAlaAsnLysTyLysArgLeuAlaAspAspLeuTyLysArgTyLysAspValAspThrGlnMet 544
DB 1349 GTGCCAATAAGTACCGCTTGCAGATGATCTTACCAGATATGATGTGATACCCAGATGT 1408
QY 544 rpThrIleLeuLysAspSerArgPheArgTyLysLeuHisThrAlaValIleValSerG 564
DB 1409 GGACCATTTCTTAGGACAGCGGATTTTCCGTTACTTGCACACAGCTGTGATAGTAGTG 1468
QY 564 lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584
DB 1469 GAACCATGCTGGTGTGTTGGGGGAAACACACACAATGACATCTATGAGCATGGCGCCA 1528
QY 584 ysCysPheSerSerAspPheMetAlaTyLysAspIleAlaCysAspArgTrpSerValLeuP 604
DB 1529 AATGCTTCTCTTCAGATTTTCATGGCCTATGACATGCTGCTGACCGCTGTGCTGCTTC 1588
QY 604 roArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaValLeuHisAsn 623
DB 1589 CCAGACCTGATTC-CACCATGATGTCAACAGATTGGCCATTTCCAGCAGTCTTACACAAC 1647
QY 624 SerThrMetTyLysValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643
DB 1648 AGCACCATGATGTGTGCTGGTGTTCATAGTCTCTCTCCTCAGCGACATCTCGGTGATTC 1707
QY 644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663
DB 1708 ACCTCGAACAAGTGTGATGGCATCGAGTGAAGCCGCTTGTGTAGCAGCAGGACCTGGT 1767
QY 664 lIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683
DB 1768 ATTGCGTGTGTGGAAACACAGGCTCTCTCAGTGTATCTCTGCGGCGCTGGCAACTGAT 1827
QY 684 GluGlnGluGlyLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703
DB 1828 GAACAAGAAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGA 1887
QY 704 CysAspGlnHisThrAspCysTyLysSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723
DB 1888 TGTGACCAAGCACACAGATTTGTACAGCTGCACAGCCAAACCAATGACTGCCACTGGTGC 1947
QY 724 AsnAspHisCysValProArgAsnHisSerCysSerGluGlnIleSerIlePheArg 743
DB 1948 AATGACCATTTGTGCCCCAGAACACACAGCTGTCTAGAAGGCCAGAGTCTCCATTTTAGG 2007
QY 744 TyrGluAsnCysProLysAspAsnProMetTyLysTyLysCysAsnLysLysThrSerCysArg 763
DB 2008 TATCAGATTTGCCCCAGAGGATTAACCTATGTACTACTGTAAACAAGAGACAGCTGCAGG 2067
QY 764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783
DB 2068 AGCTGTGCTCGACCAAGAACTGCTGAGTGGAGCCCGGAATCAGGAGTGCATTTGCCCTG 2127
```



QY 784 ProGluAsnIleCysGlyIleGlyTyrHisLeuValGlyAsnSerCysLeuIleThr 803  
 DB 2128 CCCGAAAATATCTGTGGCATGTGGCTGGCAATTTGGTTGGAATCTCAATTTTGAATTTACT 2187  
 QY 804 ThrAlaIysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823  
 DB 2188 ACTGCCAAGGAGAAATATGACAAATGCTAAATTTGTTCTGTAGGAACCCACAATGCCCTTTG 2247  
 QY 824 AlaSerLeuThrThrGlnIleLysValGluPheValLeuLysGlnLeuArgIleMetGln 843  
 DB 2248 GCTTCTCTTACAAACCCAGAGAGAGAGTAGAATTTGCTTTAAGCAGCTGCGAATAATGCAG 2307  
 QY 844 SerSerGlnSerMetSerLysLeuThrLeuThrProTyrValGlyLeuArgLysIleAsn 863  
 DB 2308 TCATCTCAGAGCATGTCCAGCTCACCTTAACCCCATGGTGGCTTCGGGAATCAAT 2367  
 QY 864 ValSerTyrTyrCysTyrGluAspMetSerProPheThrAsnSerLeuLeuGlnTyrMet 883  
 DB 2368 GTGTCTCTACTGTGCTGGGAAGATATGTCCCATTTTACAAATAGTTTACTACAGTGGATG 2427  
 QY 884 ProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGly 903  
 DB 2428 CCCTGTGAGCCCGAGTGTGATGCTGTGATTTCTGTGGAAATTTTATCAGAACCCAGTACTGGGG 2487  
 QY 904 LeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsn 923  
 DB 2488 CTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTGTAGTGTCTGTGAAGGCTGCAAC 2547  
 QY 924 HisSerAlaIysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943  
 DB 2548 CACAGTGTCTAAGCAGTGGCCGACACCATGTGGCTTGGACAGCATGTGGAGATTGACCC 2607  
 QY 944 SerGlySerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAspSerAsnAla 963  
 DB 2608 AGCGGCGAGCTCTGAGTGCATGTGTGGCAGCAACATGAACAGTGTGTGACCTCCAAATGCC 2667  
 QY 964 TyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSerThrCysPro 983  
 DB 2668 TATGTGCGCTCTCTTCTGCGCAGTGTGTGAATGTGTATGATGATGACGACCTGCCGCC 2727  
 QY 984 ProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGly 1003  
 DB 2728 CTTGAAAATTTGTCAGGCTACTGTACTGTAGTCTGTGTGGACCAACAGGCTGTGGC 2787  
 QY 1004 TrrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGly 1023  
 DB 2788 TGGTGTACTGATCCAGCAATATCTGGCAAGGGAATGATAGAGGTTCTCTATAAAGGA 2847  
 QY 1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043  
 DB 2848 CCAGTGAAGATGCTTGGCAAGCCCTCAGAGAAATTTCTATCCAGGCCCTGTCTCAAT 2907  
 QY 1044 SerSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCysProAlaCys 1063  
 DB 2908 TCAGCATGTGTCTAGAGCAGCAGACATACAACATGCTTCTTCTTCTTCTTCTTCTTCTTCT 2967  
 QY 1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083  
 DB 2968 CAATGCAACGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAACCTG 3027  
 QY 1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103  
 DB 3028 ACCACAGGCAAGCACTGCGAGACTGCATATCTGGCTTCTACGGTGTATCCCAATGGA 3087  
 QY 1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123  
 DB 3088 GGGAAATGTGAGCATGTGCAATGGCAGCGCTCTCTGTGCAACCAACCAACGCGGC 3147  
 QY 1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143  
 DB 3148 AAGTGTCTTGTGACCAACCAAGGCGTCAAGGGGAGCAGTGTGAGGTAGAA 3207  
 QY 1144 AsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyr 1163

DB 3208 AATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTATATCTCTTCTTACTACTAT 3267  
 QY 1164 GlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIleAsnPheVal 1183  
 DB 3268 CAGTTCACCTTTTGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTG 3327  
 QY 1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203  
 DB 3328 GCTACTCTCTGACGACAAACACAGGATTTGGACATGTTTATCATCAATGCTCCCAAGATTTC 3387  
 QY 1204 AsnLeuAsnIleThrTyrAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluMet 1223  
 DB 3388 AACCTCAACATCACCTGGCTGCGAGTTTCTAGCTGGAAACCCAGGCTGGAGAAGAGATG 3447  
 QY 1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243  
 DB 3448 CCTGTTGTTTCAAAACCAACCATTAAGCAGTCAAGATAGTTTCTCTAATGAGAAGTTT 3507  
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTyrPro 1263  
 DB 3508 GATTTTCGCAACACCCCAATATACATTTCTTTGTTTATGTCAGTAATTTCACTGCGCC 3567  
 QY 1264 IleLysIleGln 1267  
 DB 3568 ATCAAAATTCAG 3579

RESULT 5  
 AK000356 2578 bp mRNA linear PRI 13-SEP-2003  
 LOCUS Homo sapiens cDNA FLJ20349 fis, clone HEP13957, highly similar to  
 DEFINITION AB011120 Homo sapiens mRNA for KIAA0548 protein.  
 ACCESSION AK000356  
 VERSION AK000356.1 GI:7020387  
 KEYWORDS oligo capping; fis (full insert sequence).  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 REFERENCE 1  
 AUTHORS Kawakami,T., Noguchi,S., Itoh,T., Shigeta,K., Senba,T.,  
 Matsumura,K., Nakajima,Y., Mizuno,T., Morinaga,M., Ota,T.,  
 Suzuki,Y., Obayashi,M., Nishi,T., Shibahara,T., Tanaka,T.,  
 Nakamura,Y., Isogai,T. and Sugano,S.  
 NEDO human cDNA sequencing project  
 Unpublished  
 2 (bases 1 to 2578)  
 SUGANO,S., SUZUKI,Y., OTA,T., OYASHI,M., NISHI,T., ISOGAI,T.,  
 SHIBAHARA,T., TANAKA,T. and NAKAMURA,Y.  
 Direct Submission  
 Submitted (15-FEB-2000) Sumio Sugano, Institute of Medical Science,  
 University of Tokyo, Department of Virology, Shirokane-dai, 4-6-1,  
 Minato-ku, Tokyo 108-8639, Japan (E-mail:fldn@ims.u-tokyo.ac.jp,  
 Tel:81-3-5449-5286, Fax:81-3-5449-5416)  
 NEDO human cDNA sequencing project supported by Ministry of  
 International Trade and Industry of Japan; cDNA full insert  
 sequencing; Research Association for Biotechnology; cDNA library  
 construction. 5'- & 3'-end one pass sequencing; Department of  
 Virology and Human Genome Center, Institute of Medical Science,  
 University of Tokyo (partly supported by Science and Technology  
 Agency).  
 FEATURES  
 source Location/Qualifiers  
 1. 2578  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="HEP13957"  
 /cell\_line="HepG2"  
 /cell\_type="hepatoma"  
 /clone\_lib="HEP"  
 /note="cloning vector pME18SFL3"  
 misc\_feature 1. 2578  
 /note="highly similar to AB011120 Homo sapiens mRNA for



KIAA0548 protein"

ORIGIN

Alignment Scores:

Pred. No.: 0 Length: 2578  
Score: 564.00 Matches: 564  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 39.47% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AK000355 (1-2578)

QY	866	TyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSer	885
DB	3	TACTGTGCTGGGAAGATATGCCCATTTACAAATAGTTTACTACAGTGGATGCCGTCT	62
QY	886	GluProSerAspAlaGlyPheCysGlyLeuLeuSerGluProSerThrArgGlyLeuLys	905
DB	63	GAGCCAGTCATGCTGGATCTGTGGAAATTTATCAGAACCCAGTACTCGGGGACTGAAG	122
QY	906	AlaAlaThrCysLeuAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSer	925
DB	123	GCTGCAACCTGCATCAACCACTCAATGCTAGTGTCTGTGAAAGCCCTGCAAAACACAGT	182
QY	926	AlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGly	945
DB	183	GCTAAGCAGTCCGGACACCATGCTGCTTGGAGCAGCATGTGGAGATTGCAACGCGGC	242
QY	946	SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTrpVal	965
DB	243	AGCTCTGAGTGCATGTGGTGACCAACATGAAGCAGTGTGTGACTCCCAATGCCATGTG	302
QY	966	AlaSerPheProPheGlyGlnCysMetGluTrpTrpThrMetSerThrCysProProGlu	985
DB	303	GCCTCTCTCCCTTTGGCCAGTGTATGGAATGCTATACGATGAGCACCTGCCCTCCCTGAA	362
QY	986	AsnCysSerGlyTrpCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCys	1005
DB	363	AAATTGTTAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGGT	422
QY	1006	ThrAspProSerAsnThrGlyLysGlyLysCysLysGluGlySerTrpLysGlyProVal	1025
DB	423	ACTGATCCCAAGTACTGGCAAGGAAATGCATAGAGGGTTCCTATAAAGACCACTG	482
QY	1026	LysMetProSerGlnAlaProThrGlyAsnPheTrpProGlnProLeuAsnSerSer	1045
DB	483	AAGATGCCCTTCGACAGCCCTACAGGAATTTCTATCCACAGCCCTCTCAATTCACAGC	542
QY	1046	MetCysLeuGluAspSerArgTrpAsnTrpSerPheIleHisCysProAlaCysGlnCys	1065
DB	543	ATGTGTCTAGAGCAGACAGATACAACTGGTCTTTTCATTTCACTGTCCAGCTTGCCAAATGC	602
QY	1066	AsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThr	1085
DB	603	AACGGCCACAGTAATATGCATCAATCAGACATCTGTGAGAAATGTGAGAACCTGCACCA	662
QY	1086	GlyLysHisCysGluThrCysIleSerGlyPheTrpGlyAspProThrAsnGlyLysLys	1105
DB	663	GCCACAGCATCGGAGACTGCATATCTGGCTTCTACGGTGATCCCAACCAATGAGGGAAA	722
QY	1106	CysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCys	1125
DB	723	TGTCAGCCATGCAAGTCAATGGCAGCGGTCTGTGTGCAACCAACACAGCGGCAAGTGC	782
QY	1126	PheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuLeuGluAsnArg	1145
DB	783	TTCTGCACCAACCAAGGGCTCAAGGGGGACGAGTGCACATATGAGGTAGAAATCGA	842
QY	1146	TyrGlnGlyAsnProLeuArgGlyThrCysTrpTrpThrLeuLeuIleAspTrpGlnPhe	1165
DB	843	TACCAAGGAACCCCTCTCAGAGGAACATGTTATTACTCTTCTTATTGACTATCAGTTC	902

QY	1166	ThrPheSerLeuSerGlnGluAspArgTrpTrpThrAlaIleAsnPheValAlaThr	1185
DB	903	ACCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGTACT	962
QY	1186	ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu	1205
DB	963	CTTGACGAAACAAACAGGGATTTGGACATGTTTATCATATGCCCTCCAGAAATTTCAACCTC	1022
QY	1206	AsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProVal	1225
DB	1023	AACATCACTGGGCTGCCAGTTCTTCAGCTGGAACCCAGGCTGGAGAAGAGATGCCTGTT	1082
QY	1226	ValSerLysThrAsnIleLysGluTrpLysAspSerPheSerAsnGluLysPheAspPhe	1245
DB	1083	GTITTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAATGAGAAGTTTGAATTT	1142
QY	1246	ArgAsnHisProAsnIleThrPhePheValTrpValSerAsnPheThrTrpProIleLys	1265
DB	1143	CGCAACCAACCAATATCACTTTCTTTGTTATGTCAAGTAATTCACCTGCCCATCAAA	1202
QY	1266	IleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThr	1285
DB	1203	ATTTCAGATTCCTCTCTCAGCACAGCAATTTTATGGACCTGTACAGTTCTTCGTGACT	1262
QY	1286	PhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrpLysIleLysGln	1305
DB	1263	TTCTTCAGTTGTTTCTCTCTTTGCTCTGCTGGTGGCTGTGGTGGAAAGATCAACAA	1322
QY	1306	SerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSer	1325
DB	1323	AGTTGTTGGGCTCCAGACGTAGAGCAACTTCTTCGAGAGATGCAACAGATGGCCAGC	1382
QY	1326	ArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuLe	1345
DB	1383	CGTCCCTTTCCTCTGTAAATGTCCTTGGAAACAGATGAGGAGCTCTCTGATCTTATT	1442
QY	1346	GlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAsnLys	1365
DB	1443	GGGGGAGATATAAGACTGTTCCCAACCCATTCGACTGGAGCGGTGTTTGGCAACAAA	1502
QY	1366	AlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProProPro	1385
DB	1503	GCCCTCTCTCTCTGTGTGTGTGAGCTCCCTCGAGGCTGGGTGGCATCCCTCCCTCCT	1562
QY	1386	GlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIle	1405
DB	1563	GGGCAGTCAGGTCTGTGTGTGGCCAGCGCTTGTGGACATTTCTCAGCAGATGCCGATA	1622
QY	1406	ValTrpLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnPro	1425
DB	1623	GTGTACAAAGGAGAAGTCAGGAGCGGTGAGAAACCGGAAGCAGCAGCCCTGCACAGCCT	1682
QY	1426	GlyThrCysIle 1429	
DB	1683	GGGACCTGCATC 1694	
RESULT 6			
AR164815			
LOCUS			
DEFINITION			
SEQUENCE 12 from patent US 6274339.			
AR164815			
ACCESSION			
VERSION			
AR164815.1 GI:16238024			
KEYWORDS			
SOURCE			
Unknown.			
ORGANISM			
Unclassified.			
1 (bases 1 to 6370)			
REFERENCE			
Moore,K. and Negley,D.Lynn.			
METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT OF BODY			
WEIGHT DISORDERS, INCLUDING OBESITY			
Patent: US 6274339-A 12 14-AUG-2001;			
JOURNAL			
FEATURES			
Location/Qualifiers			
1..6370			
source			

/organism="unknown"		/mol_type="unassigned DNA"	
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	6370
Score:	503.00	Matches:	516
Percent Similarity:	99.81%	Conservative:	0
Best Local Similarity:	99.81%	Mismatches:	0
Query Match:	35.20%	Indels:	1
DB:	6	Gaps:	0
US-09-787-097-12 (1-1429) x AR164815 (1-6370)			
QY	914	AsnGlySerValCysGluArgProAlaAsnHisSerAlaLys-GlnCysArgThrProCy	933
DB	506	AATGGTAGTGTCTGTGAAGGCTCGAACCACACAGTGTCTAAGGCAGTGGCGGACACCATG	565
QY	933	sAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSe	953
DB	566	TGCCTTGAGACAGCATGTGGAGATTGCCACCGCGGAGCTCTGAGTGCATGTGGTGCAG	625
QY	953	rAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCy	973
DB	626	CAACATGAAGCAGTGTGGACTCCAAATGCTATGTGGCTCTCTCCCTTTTGGCCAGTG	685
QY	973	sMetGluTrpTyrThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCy	993
DB	686	TATGGAATGGTATACGATGAGCAGCTGCCCTCCCTGAAATTTGTTACGGCTACTGTACCTG	745
QY	993	sSerHisCysLeuGluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLy	1013
DB	746	TAGTCATTTGCTGGAGCAACCGAGCTGTGGCTGTACTGATCCAGCAATACTGGCAA	805
QY	1013	sGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProTh	1033
DB	806	AGGGAATGCTATAGAGGGTTCTTATAAGGACCAAGTGAAGATGGCTTCGACAGCCCTAC	865
QY	1033	rGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTy	1053
DB	866	AGGAATTTCTATCCACAGCCCTGCTCAATTCAGCATGTGTCTAGAGACAGCAGATA	925
QY	1053	rAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAs	1073
DB	926	CAACTGGTCTTTCACTCTCCAGCTTGCCTTCCCAATGCAACGGCCACAGTAATGATCAA	985
QY	1073	nGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysI	1093
DB	986	TCAGAGCATCTGTGAGAAAGTGTGAGAACCTGACACAGGCAAGCACTGGCAGACCTGCAT	1045
QY	1093	eSerGlyPheTyrGlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGl	1113
DB	1046	ATCTGGCTTCTACGGTATCCCAATGAGGAGGAATGTGACGCATGCAAGTGCATGG	1105
QY	1113	yHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyVally	1133
DB	1106	GCACGGCTCTGTGCAACACCAACAGCGGCAAGTGTCTTCGACCAACCAAGCGCGTCAA	1165
QY	1133	sGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGl	1153
DB	1166	GGGGGACGAGTGCAGCTATGTGAGTAGAATAATCGATACCAAGGAACCCCTCTCAGAG	1225
QY	1153	yThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAs	1173
DB	1226	AACATGTTATTATACCTCTTCTTATGACTATCATCTACCTTTTATGCTATATCCCGAGAGA	1285
QY	1173	pAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLe	1193
DB	1286	TGATCGCTATTACACAGCTATCAATTTTGTGGTACTCTCTGACGAACAAACAGGGATT	1345
QY	1193	uAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPh	1213
DB	1346	GGACATGTTCAATGCTCTCCAGAAATTTCAACCTCAACATCACCTGGGCTGCCAGTTT	1405
QY	1213	eSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGl	1233
DB	1406	CTCAGCTGGAACCCAGGCTGGAGAGAGATGCTCTGTTTCAAAAAACCAACATTAAAGA	1465
QY	1233	uTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPh	1253
DB	1466	GTACAAAGATAGTTTCTTAATGAGAAGTTTGAATTTTCGCAACACCAAAATATCACTTT	1525
QY	1253	ePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHi	1273
DB	1526	CTTTGTTTATGTCAAGTAATTTTCACTGGCCCATCAAAATTCAGATTGCCCTTCTCAGCA	1585
QY	1273	sSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLe	1293
DB	1586	CAGCAATTTTATGACCTGGTACAGTTCTTCTGTCGACTTTTCTTCAGTTGTTTCTCTCTTT	1645
QY	1293	uLeuLeuValAlaAlaValTyrLysIleLysGlnSerCysTrpAlaSerArgArgAr	1313
DB	1646	GCTCCTGGTGGCTGCTGTGGTTTGAAGATCAAAACAAAGTTGTGTGGGCTTCCAGACGTAG	1705
QY	1313	gGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnVa	1333
DB	1706	AGAGCAACTTCTTCGAGAGATGCAACAGATGGCCGCGCTGCCCTTTCCTCTCTGTAATGT	1765
QY	1333	lAlaLeuGluThrAspGluGluProProAspLeuIleGlyLysSerIleLysThrValPr	1353
DB	1766	CGCTTGGAAACAGATGAGGAGCTCTCTGATCTTATTTGGGGGAGTATAAAGACTGTTC	1825
QY	1353	oLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheVa	1373
DB	1826	CAAAACCATTCACCTGGAGCCGTGTTTGGCAACAAAGCCGCTGCTCTCTCTGTGTTGT	1885
QY	1373	lArgLeuProArgGlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAl	1393
DB	1886	GAGGCTCCCTCGAGCCCTGGGTGGCATCCCTCTCTCTGGGAGTCAGGTCTTCTGTGGC	1945
QY	1393	aSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAl	1413
DB	1946	CAGCGCCCTGGTGGACATTTCTCAGCAGATGCCATAGTGTACAGGAGAGATCAGGAGC	2005
QY	1413	aValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle	1429
DB	2006	CGTGAAACCCGGAAGCAGCAGCCCTGACAGCCCTGACAGCCTGGGACCTGCATC	2054
RESULT 7			
AR164818			
LOCUS			
DEFINITION			
SEQUENCE			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
1..2625			
/organism="unknown"			
/mol_type="unassigned DNA"			
ORIGIN			
Alignment Scores:			
Pred. No.:	0	Length:	2625
Score:	496.00	Matches:	696
Percent Similarity:	99.43%	Conservative:	0
Best Local Similarity:	99.43%	Mismatches:	2
Query Match:	34.71%	Indels:	4
DB:	6	Gaps:	0

US-09-787-097-12 (1-1429) x AR164818 (1-2625)	
QY	87 AlaAlaValSerGlySerAlaAlaGluAlaLysGluCysAspArgProCysVal 106
Db	256 GGGGGGGGGTTCGGCTCGCCGCGAGCCGAGGCAAGAAATGACCGCGCGTGTGC 315
QY	107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTyrValGly 126
Db	316 AACGGCGGTTCGGTGCACCTGGCACCGGCGAGTGGCTCGCCCGCGGTGGTGGGC 375
QY	127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
Db	376 GAGCAATGCGACACTGCGGGGGCGGCTTCAGACTAACTGGATCTCTCGGGTTTGACA 435
QY	147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuLeuLeuGlnPro 166
Db	436 GATGGACCTGGAAATATATAAATACAAACGAAGTGCACGTGCTCATTAAGGACGACA 495
QY	167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
Db	496 AATAGAATAAATGAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTTGGACCATTTA 555
QY	187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuLe 206
Db	556 TATGTTTATGATGGGAGCTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCCCTCAT 615
QY	207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
Db	616 GTTCTCTGAGAGATGGCAATGAGACTGTCCTCTGAGGTGTTGCCACATCAGTTATGCC 675
QY	227 LeuLeuHisPhePheSerAspAlaLalaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
Db	676 TTGCTGCATTTTTTATGATGCTGCTTATAATTTGACTGGATTTAATATTACTTTACAGT 735
QY	247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266
Db	736 TTGATATGTGTCCAAATAACTGCTCAGCCCGAGGAGGTGAAGATCAGTAATAGACAGC 795
QY	267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHI 286
Db	796 GA-TACTGTTGAATGTAATGTTCTGAAAACTGGAAAGGTGAAGCATGTGACATTCCTCA 854
QY	286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306
Db	855 CTGTACAGACAACTGTGGTTCCTCATCGAGCATCTGCAATCAAGTGAATGTCAGAGG 914
QY	306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326
Db	915 ATGCTCTGCTTCTCAGACTGGCAGGGTCTCTGATGTTTCAGTTCCTGTACCACTAACCA 974
QY	326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisGlyAl 346
Db	975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGCAGAGCATCTCATAAAGC 1034
QY	346 aValValAsnGlyAsnIleMetTrpValValGlyTyrMetPheAsnHisSerAspTyr 366
Db	1035 TGTGTCATATGGAACATTTATGTGGGTGTTGTGAGGATATATGTTCAACCACTCAGATTA 1094
QY	366 xAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
Db	1095 TAAATGGTTCAGCGTATGACCTGCTTCTAGGGAGTGGCTTCACCTAAACCGTCTGT 1154
QY	386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
Db	1155 GAACAAATGGTGTGTAGATATGCTCATTTTGGCATTATTAACAGGATAAAATTTTACAT 1214
QY	406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
Db	1215 GTATGGAGGAAATTTGATTC-AACTGGGAATGTGACCAATCAGTTGAGAGTTTTTCA 1273
QY	426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446

RESULT 8  
AB011120  
LOCUS

5632 bp mRNA linear PRI 10-APR-1998

DEFINITION Homo sapiens mRNA for KIAA0548 protein, partial cds.  
 ACCESSION AB011120  
 VERSION AB011120.1 GI:3043619  
 KEYWORDS KIAA0548 protein.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (sites)  
 AUTHORS Nagase,T., Ishikawa,K., Miyajima,N., Tanaka,A., Kotani,H., Nomura,N. and Ohara,O.  
 TITLE Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro  
 JOURNAL DNA Res. 5 (1), 31-39 (1998)  
 MEDLINE 96290545  
 PUBMED 9628581  
 REFERENCE 2 (bases 1 to 5632)  
 AUTHORS Ohara,O., Nagase,T. and Ishikawa,K.  
 TITLE Direct Submission  
 JOURNAL Submitted (13-FEB-1998) Osamu Ohara, Kazusa DNA Research Institute, DNA Technology, 1532-3 Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:cdmainfo@kazusa.or.jp, Tel:+81-438-52-3913, Fax:+81-438-52-3914)

## FEATURES

Location/Qualifiers  
 1..5632  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="HH0769"  
 /sex="male"  
 /tissue type="brain"  
 /clone\_lib="pBluescriptII SK plus"

## gene

1..5632  
 /gene="KIAA0548"

## CDS

<1..1361  
 /gene="KIAA0548"  
 /codon\_start=3  
 /product="KIAA0548 protein"  
 /protein\_id="BA25474.1"  
 /db\_xref="GI:3043620"

/translation="TMSTCPENCSCGYCTCSHCLPQCGWCCTDSNTGKCKTIGSY  
 KGPVMSQAPTFYFQPLNLSMCLEDSRYNWSFIHCPACCCNGHSCINQSI  
 CENLTGKHCETICISGYDPTNGKQCPCKNGHSLCNTNTGKCTCTGKVGDEC  
 QLCVENRYQGNPLRGCTCYVLLIDYFTFSLQSDDRXYTAINEFVATPDEGNRLDM  
 FINASPNFLNTWAASRPSAGTQAGEMPVSVKTNKEVDSFNEKEDFRNHNPTF  
 FVYSNFTWPIKTIQAFSOHSNFMOLVOFVTFSCFLSLVLAVALPEKIKOSWASR  
 REQLREQQWASRFPASVNVVALETDDEPDLIGSIKTVKPIALEPCFGNKA  
 SVFVRUPRGLGGIPPPGQSLAVASALVDISOOMPFIYVKEKSGAVNRKQPPAOPGT  
 CI"

## ORIGIN

Alignment Scores:  
 Pred. No.: 0 Length: 5632  
 Score: 452.00 Matches: 452  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 31.63% Indels: 0  
 DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AB011120 (1-5632)

Qy 978 ThrMetSerThrCysProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeu 997  
 Db 3 ACGATGAGCACCTGCCCCCTGAAATTTGTCAGGCTACTGTACCTGTAGTCAATGCTTG 62  
 Qy 998 GluGlnProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCys 1017  
 Db 63 GAGCACCAGGCTGTGGCTGGTGTACTGATCCAGCAATACTGGCAAGGGAATGCATA 122  
 Qy 1018 GluGlySerTrpLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyr 1037  
 Db 123 GAGGTTCTTATAAGACCAAGTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTAT 182

Qy 1038 ProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPhe 1057  
 Db 183 CCACAGCCCTGTCTCAATTCAGCATGTGTCTAGAGCAGCAGATACAACTGGTCTTTTC 242  
 Qy 1058 IleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCys 1077  
 Db 243 ATTCACTGTTCAGCTTGCCTCAATGCAACGGCCACAGTAATGCATCAATCAGACATCTGT 302  
 Qy 1078 GluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyr 1097  
 Db 303 GAGAAGTGTGAGAACCTGACACAGGCAAGCACTGGCAGACCTGCATATCTGCTTTCTAC 362  
 Qy 1098 GlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeu 1117  
 Db 363 GGTGATCCCACCAATGAGGGAATGTTCAGCAATGCAAGTGCAATGGCAGCGTCTCTG 422  
 Qy 1118 CysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCys 1137  
 Db 423 TGCAACACCAACACGGCGCAAGTGTCTTCGACCAACCAAGGGCGTCAAGGGGACGAGTGC 482  
 Qy 1138 GlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyr 1157  
 Db 483 CAGCTATGTGAGGTAGAAATTCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTAT 542  
 Qy 1158 ThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyr 1177  
 Db 543 ACTCTCTTATTGACATCAGTTACCTTTAGTCTATCCAGGAGATGATGCTATTAC 602  
 Qy 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197  
 Db 603 ACAGCTATCAATTTTGTGGCTACTCTCTGACGAAACAAACAGGATTTGGACATGTTTATC 662  
 Qy 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTyrAlaAlaSerPheSerAlaGlyThr 1217  
 Db 663 ATGCTCTCCAGAAATTTCAACCTCAACATCCTGGCTGCCAGTTTCTCAGCTGGNACC 722  
 Qy 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237  
 Db 723 CAGCTCGAGAGAGATGCTGTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGT 782  
 Qy 1238 PheSerAsnGluLysPheAspAspArgAsnHisProAsnIleThrPhePheValTyrVal 1257  
 Db 783 TTCTCTAATGAGAAGTTTGATTTTCGACACCCCAAAATACACTTCTTTGTTTATGTC 842  
 Qy 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277  
 Db 843 AGTAATTTCACTGGCCCATCAAAATTCAGATTGCCCTTCTCTCAGCACAGCAATTTATG 902  
 Qy 1278 AspLeuValGlnPhePheValThrPhePheSerCysPheIleuSerLeuLeuValAla 1297  
 Db 903 GACCTGTGACAGTTCTTCGTGACTTTCTCAGTTGTTCCTCTCTTTCCTCTGGTGGCT 962  
 Qy 1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeu 1317  
 Db 963 GCTGTGTTTGGAGATCAACAAAGTTGTGGCCCTCCAGACGTAGAGAGCAACTTCTT 1022  
 Qy 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337  
 Db 1023 CGAGAGATGCAACAGATGGCCCGCTCCCTTTGCCCTCTGTAATGTGCGCTTGGAAACA 1082  
 Qy 1338 AspGluGluProProAspLeuIleGlySerIleLysThrValProLysProIleAla 1357  
 Db 1083 GATGAGAGCTCTCTGATCTTATTGCGGGAGTATAAGACTGTTCCTCAAAACCAATGCA 1142  
 Qy 1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377  
 Db 1143 CTGAGCGGTGTTTGGCAACAAAGCCGCTGCTCTCTCTGTGTGTGTGTGAGGCTCCCTCGA 1202  
 Qy 1378 GlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397  
 Db 1203 GGCCTGGTGGCATCCCTCTCTCTGGGCAGTCAGTCTTCTGTGGCCAGCGGCTGGTG 1262

QY 1398 ApilSerGlnMetProileValTyrLysGluLysSerGlyAlaValArgAsnArg 1417  
 |||||  
 Db 1263 GACATTTCTCAGCAGATGCCGATAGTGTACAAAGGAGAGTCCAGGAGCGGTGAGAAACCGG 1322  
 |||||  
 QY 1418 LysGlnGlnProProAlaGlnProGlyThrCysIle 1429  
 |||||  
 Db 1323 AAGCAGCAGCCCTGCACAGCTGGGACCTGCATC 1358  
 |||||

## RESULT 9

AF531101  
 LOCUS 4493 bp mRNA linear NAM 19-AUG-2002  
 DEFINITION Bos taurus attractin mRNA, complete cds.  
 ACCESSION AF531101  
 VERSION AF531101.1 GI:22297307

KEYWORDS Bos taurus (cow)  
 SOURCE Bos taurus  
 ORGANISM Bos taurus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Butheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 Bovidae; Bovinae; Bos.

REFERENCE 1 (bases 1 to 4493)  
 Graphodatskaya,D., Joerg,H. and Stranzinger,G.  
 Direct Submission  
 Submitted (18-JUL-2002) Animal Sciences, ETH, Tannenstrasse 1,  
 Zurich, CH 8092, Switzerland

FEATURES  
 Location/Qualifiers  
 1..4493

/organism="Bos taurus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9913"  
 /chromosome="13"  
 52..4299

/note="transmembrane form; mahogany protein"  
 /codon\_start=1  
 /product="attractin"  
 /protein\_id="AA095445.1"  
 /db\_xref="GI:22297308"

/translation="MVAAATEARLRTVTVPAPTSRNGRRCDWDATRAPRPLRTGL  
 RLPRWRRLILLPLPILLPAAEAAAASVSSAAAEKEDRCPVNGRCNPGT  
 GCGCPAGVQCQCHGRFLTSGSGFITDGNKYKTKTWLIRGQNRIMRLR  
 FNHATGSDWHLVYDGDLSIAPLVAAPSGILVPERDNETVPEVATSGVALLHP  
 SDAAYNLGPNITYNFDMPCNCSGRGCKISNSNTVQCSSEWKEACDIPHCVN  
 NGCPHRRGICNSDVRGSCFSEWGPCSVFPANQSFWTREYSIDLKPRASHKAV  
 VNGNIVVGGVFMFNSYNMVLYDLASREWLNRNSVNSVVRVYGHSLALYQDKY  
 YFGKIDSTGNVITRFLVHINESVLLSPKAEQYAVVGHSAHIVTLKNGRVMLV  
 IFGCHPLYSVVEQYEDLKNITLITQALVQCGYSHSVYDDRTKALYIHGGYK  
 AFSANKYRLADLVYDVDTOMWTILKDSRFRLYLTAVIVSGTMLVFGNTHDSTM  
 SHGACFSDFMAYDIDCDRWSVLPRDLHHDVAFEGHSAVLNRYVFGFNSLLI  
 SDILVFTSECEAHQSEAEALAGPVRGCMQDTSQCSWELAPEAGEKIKSCFSK  
 RIFDRDCQUTDCSTANTNGCOWCHDNCVPTNHSCTGCIISIFKYDHPKDNPMY  
 YNKKTSRSCALDQNCQWEPNQCIALPENICGIGWHLVNSCLKITTAETDYNA  
 KLSCRNHAFLASLTQKKVFLKQLRIMQSSQMSKFTLTPWVGLRKINVSYWCWE  
 DMSPTNSLQWMPSEPDPADGFCGLISPESTRGLKAATINPLNGSVCCERPAHSAQ  
 CRTPCALRTACCECTSGSECCWCSNMKQCVDSNAYVASFPFGCMEWYTWSSQPPEN  
 CSYCTCSCHLQPCGCHCTQPSNTGKCKLEGSKYKVPKSPQGTGNSYPOILLAS  
 AGEENPVVSKTKIKYKDSFGNEKDFRNHPNITFFVYSNFTWPKIKIYAFSGHSF  
 MDLVQFFVTFSCFLSLLLVAAVVMKIKQSCWAGRRRQLREMOQMASRPFASVNA  
 LETDEPPDLIGSLIKTVKPIALEPCFGNKAALVLSVFLRPLRGLGGIPPPQSGSLAV  
 ASALVDISQMPVYVYKESGAVRNKQPPAPQPTCI"

## ORIGIN

Alignment Scores:  
 Pred. No.: 9,76e-232 Length: 4493  
 Score: 252.00 Matches: 574  
 Percent Similarity: 97.29% Conservative: 0  
 Best Local Similarity: 97.29% Mismatches: 8  
 Query Match: 17.63% Indels: 16  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AF531101 (1-4493)

QY 823 LeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMet 842  
 |||||  
 Db 2476 TTGGCTTCTCTTACAACCTCAGAGAGAGTAAATTTGCTCTTAAGCAGCTCGCAATCATG 2535  
 |||||  
 QY 843 GlnSerSerGlnSerMetSerLysLeu-ThrLeuThrProTyrValGlyLeuArgIle 862  
 |||||  
 Db 2536 CAGTCATCAGAGAGATGTCAGATT-CACCTTAACCCCATGGTGGCTTCGGAAGAT 2594  
 |||||  
 QY 862 eAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsnSerLeu-LeuGlnT 882  
 |||||  
 Db 2595 CAATGTCTCTATTGGTGTGGGAGATATGTCCCCGTTCAAAATAGTTC-ACTACAGT 2653  
 |||||  
 QY 882 xPmetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrA 902  
 |||||  
 Db 2654 GGATGCCATCTGAGCCCGAGTATGCTGATCTGTGGGATCTTGTCAAGACCCAGTACTC 2713  
 |||||  
 QY 902 rgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProA 922  
 |||||  
 Db 2714 GGGGCTTTAAAGGCTGCAACCTGCATCAATCCACTCAATGCTAGTGTCTGTGAAAGGCTG 2773  
 |||||  
 QY 922 laAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys-GlyAsp 941  
 |||||  
 Db 2774 CTAACACAGTGGCCAGCAGTGTGGACCCCTTGTGCTTGGCAGACGCTGTGGGAG- 2832  
 |||||  
 QY 942 CysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAspSer 961  
 |||||  
 Db 2833 TGTACACAGTGGCAGCTCGGAGTGCATGTGTGTCAGCAGACATGAAGCAGTGGTGGACTCC 2892  
 |||||  
 QY 962 AsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSerThr 981  
 |||||  
 Db 2893 AATGCCCTACGTGGCTCTTCCCTTTTGGCCAGTGTATGGAGTGGTATACCATGAGCAG- 2951  
 |||||  
 QY 982 -CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnProGln 1001  
 |||||  
 Db 2952 CTGGCCCCCTTGAAATTTGTTCAGGCTACTGTACTGTAGTCACTGTCTGGAGCAACGAG 3011  
 |||||  
 QY 1001 YCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyr 1021  
 |||||  
 Db 3012 CTGTGGCTGTGTGCTGCTGCTAGCAATACGGGCAAGGAAATGCTAGAGGGCTCTCTA 3071  
 |||||  
 QY 1021 rLysGlyProValLysMetProSerGlnAla-ProThrGlyAsnPhe-TyrProGlnPro 1040  
 |||||  
 Db 3072 TAAAGGACCAAGTGAAGTGCCTTCTCAGG-CCCTACAGAAA-TTCCTACCCACAGCCT 3129  
 |||||  
 QY 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCys 1060  
 |||||  
 Db 3130 CTTCTCAATTCAGCAGATGTCTCTGGAGGACAGCAGATATAACTGGTCTTTCATTGT 3189  
 |||||  
 QY 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGlyLysCys 1080  
 |||||  
 Db 3190 CCAGCTTGCAGTGCATGTCATGCCACAGCAATGCAATTAATCAGAGTATCTGTGAGAGTGT 3249  
 |||||  
 QY 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100  
 |||||  
 Db 3250 GAGAACTGACACGGGCAAGCACTGTGAGACCTGCAATCTGGCTTCTATGTGTATCCA 3309  
 |||||  
 QY 1101 ThrAsnGlyGlyLysCysGlnProCysLys-CysAsnGlyHisAlaSerLeuCysAsnTh 1120  
 |||||  
 Db 3310 ACCAACGGGGGAAAATGTCAACCATGCAG-GTGCAATGGACATGCGTGGCTGATGCAATAC 3368  
 |||||  
 QY 1120 rAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCy 1140  
 |||||  
 Db 3369 CAACACGGGCAAGTGTCTTCTGCACCCACCAAGGGGCGTCAAGGGGAGTGAATGTGACGTATG 3428  
 |||||  
 QY 1140 sGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuL 1160  
 |||||  
 Db 3429 TGAGGTAGAAAATCGATACCAAGGAAACCTCTCAA-AGGAACGTGTATTATACTCTTC 3487  
 |||||  
 QY 1160 eulleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaI 1180  
 |||||  
 Db 3488 TTATTGACTATCAGTTTACCTTTAGCCTATCCCGAGGAAGATGACCGCTATTACACAGCCA 3547  
 |||||





QY 942 --CysThrSerGly-SerSerGluCysMetTTPCysSerAsnMetLysGlnCysValAsp 960  
Db 2816 AGTGCACTAGCAG-CAGCTCTCAGTGCATGTGTGGTGCAGCAACATCAAGCAGTGTGTAGAC 2874  
QY 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSer 980  
Db 2875 TCCAAATGCCATATGTGGCCCTCCCTTCCTTGTGGCCAGTGTATGGAGTGTGTATACAAATGAGC 2934  
QY 981 Thr-CysProGluAAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPr 1000  
Db 2935 AG-CTGCCACCTGAAATTTGCTCTGGCTACTGTACCTGCAGCCATTGCTTGGAGCAGCC 2993  
QY 1000 oGlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySe 1020  
Db 2994 AGCGTGTGCTGTGTACTGTACTGTAGCAATATCTGGGAAGGGAATGTATCGAGGGCAG 3053  
QY 1020 rTyLysGlyProValLys-MetProSerGlnAlaPro-ThrGlyAsnPhe-TyrProGln 1039  
Db 3054 CTATAAGAGACCCAT-AAGATGCTCTACAGGCTTC-TACAGGAAATGT-CTATCCACA 3110  
QY 1039 nProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHi 1059  
Db 3111 GCGCTCTCTGAACCTAGCATGTGTTAGAGCAGCAGATACAACTGGTCTTTTCATTCA 3170  
QY 1059 sCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079  
Db 3171 CTGTCCAGCTTCCAGTGCATAGGACACAGCAAAATGTATTACACAGAGTATCTGTGAGAA 3230  
QY 1079 sCysGluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099  
Db 3231 GTGTGAGAACCTGACCACTGGCAAGCAGCTGTGAGACCTGCATATCTGGCTCTATGGTGA 3290  
QY 1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119  
Db 3291 TCCGACTAATGGAGCAAAATGTCCGCCATGTAAGTGCAACGGCGCATGCTGCTGTGCAA 3350  
QY 1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139  
Db 3351 TACCACACGGCAAGTGTCTTGACCAACGAGGTGTCAAGGAGATGAGTGCAGCT 3410  
QY 1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159  
Db 3411 ATGTGAGGTAGAAAATCGATACCAAGGAACCCCTCTCAA-AGGAACATGTTATTATGCC 3469  
QY 1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnLeuAspArgTyrTyrThrA 1179  
Db 3470 TTCTCATTCACATATCAGTTTCACCTTTTAGCCTGTCCAGGAACAGCATGCTATTATACAG 3529  
QY 1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199  
Db 3530 CCATCAATTTTGTGGCCACTCTCTGATGAACAAAACAGGGATTTGGACATGTTTATCAATG 3589  
QY 1199 laSerLysAsnPheAsnLeuAsnIleThrTyrAlaAla-SerPheSerAlaGlyThrGln 1218  
Db 3590 CCTCAAAAACCTTCAACCTCAACATCACCTGGGCTAC-CAGCTTCTCAGCTGGAACCCAG 3648  
QY 1219 AlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPhe 1238  
Db 3649 GCAGGAGAGAAATGCTGTGTTTCAAAAACCAACATTAAGAGTACAAAGATAGCTTC 3708  
QY 1239 SerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSer 1258  
Db 3709 TCTAAACGAGAAATTTGATTTTTCGCAACCATCCAAAACATTACTTCTTTTATGTCAGT 3768  
QY 1259 AsnPheThrTyrProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278  
Db 3769 AATTTCCACCTGGCCCATCAAAATTCAGATTCGCTTCTCCAGACAGCAATTTTATGGAC 3828  
QY 1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAla 1298  
Db 3829 CTGTGACAGTCTTCTGTGACTTCTTCTCAGCTGTTTCTCTCTCTCTCTCTCTCTCTCTCT 3888  
QY 1299 valValTrpLysIleLysGlnSerCysTyrTrpAlaSerArgArgGluGlnLeuLeuArg 1318

Db 3889 GTGGTTTGGGAAGATCAAGCAGAGCTGTGGGCACTCCAGAGGAGGAGCAACTTCTTCGG 3948  
QY 1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp 1338  
Db 3949 GAGATGACAGATGGCCAGCCGTCCTTTCCTTCTGTAATGTTGCTTGGCTTGGAAACAGAT 4008  
QY 1339 GluGluProProAspLeuIleGlySerIleLysThrValProLysProIleAlaLeu 1358  
Db 4009 GAAGAGCCTCTGATCTATTGCGGGAAGTATAAGACTGTCCCAAGCCATTGCCCCTG 4068  
QY 1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly 1378  
Db 4069 GAGCCCTGCTTTGGTAAACAGCTGCAGTCTCTCTGTTATTGTAGGCTCCCTTCGAGGA 4128  
QY 1379 LeuGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAsp 1398  
Db 4129 CTGGTGGAAATCCCTCTCTCTGCTCAGTCAGTCAAGTCTGCTGCGCAGTGCCTGGTGAC 4188  
QY 1399 IleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLys 1418  
Db 4189 ATTTCTCAACAGATGCCGATAGTATACAAAGAGTCAGAGCGGTAAAGAAACCGGAAG 4248  
QY 1419 GlnGlnProProAlaGlnProGlyThrCysIle 1429  
Db 4249 CAGCAGCCCCCTGCACAGCCTGGACCTGCATT 4281  
RESULT 11  
AF119821 4313 bp mRNA linear ROD 16-APR-1999  
LOCUS Mus musculus attractin (Mgca) mRNA, complete cds.  
DEFINITION AF119821  
ACCESSION AF119821  
VERSION AF119821.1 GI:4585306  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
1 (bases 1 to 4313)  
AUTHORS Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,  
Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.  
TITLE The mouse mahogany locus encodes a transmembrane form of human  
attractin  
JOURNAL Nature 398 (6723), 152-156 (1999)  
MEDLINE 99184160  
PubMed 10086356  
REFERENCE 2 (bases 1 to 4313)  
AUTHORS Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,  
Schlossman,S.F., Duke-Cohan,J.S. and Barsh,G.S.  
TITLE Direct Submission  
JOURNAL Submitted (12-JAN-1999) HMI, Stanford, Beckman Center B271A,  
Stanford, CA 94305-5323, USA  
FEATURES  
source location/Qualifiers  
1..4313 /organism="Mus musculus"  
/mol\_type="mRNA"  
/db\_xref="taxon:10090"  
/chromosome="2"  
1..4313 /gene="Mgca"  
/note="mahogany candidate"  
27..4313 /gene="Mgca"  
/note="mahogany"  
/codon\_start=1 /product="attractin"  
/protein\_id="AAD25372.1"  
/db\_xref="GI:4585307"



KGBSCDI PHCTDNCGPHRGI CNASDTRGSCFPHWQPGCSIPVPANOSFWTRBEYS  
DLKLPASHKAVNGNIMVVGGMFNHSDYSMLVAYDLTSREWLPLNHSVNSVVRVY  
GSLALHDKIYMYGKIDGNTGNVNELEFVHNESSWLLTPKAKDOXAVGHSAAHI  
VTLASGRVNLVIFGHICPLIYGYISVVOEVDLEKNTWSILHTQGLVQGGYGHSSAYDD  
RTKALYHVGYSKAFKANKYRLADDLRYDVTOMTILKDSRFFRYLHTAVLSTGML  
VFGNTHNTDSHSHGAKCFSSDFMAYDIACDRSVLPRLHHDVNRFGHSALVNST  
MYVFGGNSLLSLVLFVTSQCDAHRSEAAVCAAGPIRCLWDVQSRCTSWELATE  
EOAKLSECFKSRKTLDRCHDOHTCYCTANTINDCHWCHDVPVNHSCTEGQISI  
AKYESCKPNMPCYCNKTSKSCSALDQNCONEPRNOECIALPENICGNHVLVNSC  
LKITAKENYDNALSKRNHAFSLASTSQKVELVAKOLRLMQSSQMSKLTLPWV  
GLKINVTWCHMEDMFFNTSLUOWPSPSDAGFCGLISESTEGSLKATCINPLNG  
SVGERPANHSKQRTPCALRTACGECTSSSCSNKQCVDSNAYVAFPPGQC  
MEWYTSPPENSCSYCTCSHGLEQGGWCTDPSNTGKICEGYSKYPVMPQSA  
SAGNVPQPLNSMCLDSRYNWSFHCPCQKSHKINQSCVCEKEDLTGKHG  
ETCISGFVGDPTNGGCKQPKCNHSHASLNTGKFCFTTKGVKGDCEOLCKVENRYQ  
GNPLKCTCYVTLIDYQFTFSLSQGDRVYTAINFVATPDEQNRDPMFINAKXFNL  
NITWASFPAGTGTGEVFPVSKTNIKEYKDSFSEKEDPFRNHPNITFPVYVNSTP  
IKIQIAPFSQHSNFDVJQVFFVFFSCLLSLLVAAVWKIKOSCNASRRERQLLREM  
QMASRPFASVNVVALEDEPDLIGSIKTVKPIALEPCFNGKAAVLSVFLPRGL  
GGTTPPQSGSLAVASALVDISQMPVIYVKEKSGAVNRKXQPPAQPGTCTI"

## ORIGIN

Alignment Scores:  
Pred. No.: 1,82e-187 Length: 4313  
Score: 206.00 Matches: 206  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 14.42% Indels: 0  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x AFI119821 (1-4313)

Qy 1224 ProValSerLysThrAsnIleLysGluTyLysAspSerPheSerAsnGluLysPhe 1243  
Db 3693 CTTGTTGTTTCAAAACCAACATCAAGGAATACAAAGATAGTCTTCTAATGAGAAATTT 3752  
Qy 1244 AspPheArgSerHisProAsnIleThrPhePheValTyValSerAsnPheThrTrpPro 1263  
Db 3753 GATTTCGCAACCATCAACATCATCTTTCTTTGTTATGTCAGTAATTTCTACTTGGCCCC 3812  
Qy 1264 IleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePhe 1283  
Db 3813 ATCAAAATTCAGATTGCTTCTCCAGCAGCAGCACTTTATGACCTGTGACAGTCTTCTC 3872  
Qy 1284 ValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrpLysIle 1303  
Db 3873 GTGACTTTCTTCAAGTTGTTTTCCTCTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCT 3932  
Qy 1304 LysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMet 1323  
Db 3933 AAGCAGAGCTGTTGGGATCCAGCGGAGAGAGACAACTTCTTCGGGAGATGCAACAGATG 3992  
Qy 1324 AlsSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAsp 1343  
Db 3993 GCCAGCGCCCTTGTGTTCTGTAACCGTTGCTTGTGAAACGTTGCTTGTGAAACGATGA 4052  
Qy 1344 LeuIleGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGly 1363  
Db 4053 CTATTGGGGGAGATGTAAGACCGTTCTTAAGCCATTCCTGAGCCCTGAGCCCTGCTTGT 4112  
Qy 1364 AsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyLePro 1383  
Db 4113 AACAAAGCCGAGTCTCTCTGTTATCGTAGGCTCCCTCGAGGACTGGGAGGAAATCCCT 4172  
Qy 1384 ProGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMet 1403  
Db 4173 CCTCCTGGTCAAGTCTGCTGTCGTCGAGCGGCTGCTGTCGAGCATTTCTCAGCAGAT 4232  
Qy 1404 ProIleValTyLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProAla 1423  
Db 4233 CCAATAGTGTAAGGAGAGAGTCAAGGAGCTGTAAAGAACCGGAGAGCAGCGCGCTGCA 4292  
Qy 1424 GlnProGlyThrCysIle 1429

Db 4293 CAGCCTGGAACCTGCATT 4310  
RESULT 12  
AB038387  
LOCUS Rattus norvegicus atrin mRNA for attractin, complete cds.  
DEFINITION AB038387  
ACCESSION AB038387.1 GI:12275307  
VERSION  
KEYWORDS attractin.  
SOURCE Rattus norvegicus (Norway rat)  
ORGANISM Rattus norvegicus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (sites)  
Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,  
Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.  
and Serikawa,T.  
Attractin/mahogany/zitter plays a critical role in myelination of  
the central nervous system  
Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)  
21143347  
11209055  
2 (bases 1 to 8739)  
Kuramoto,T., Serikawa,T. and Ushijima,T.  
Direct Submission  
Submitted (15-FEB-2000) Takashi Kuramoto, National Cancer Center  
Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,  
Chuo-Ku, Tokyo 104-0045, Japan [E-mail:tkuramoto@ncc.ncc.go.jp.  
Tel:81-33542-2511 (ex.4521), Fax:81-35565-1753]  
Location/Qualifiers  
FEATURES  
source  
1..8739  
/organism="Rattus norvegicus"  
/mol\_type="mRNA"  
/strain="WTC"  
/db\_xref="taxon:10116"  
1..8739  
/gene="atin"  
61..4359  
/gene="atin"  
/codon\_start=1  
/product="attractin"  
/protein\_id="BAB21017.1"  
/db\_xref="GI:12275308"  
/translation="MVAAAASAEATEARLGVTATAAPAGWKEORHRPCAATGAMRP  
WPRAGLCIPRVLSRALSPPELLELPLFLSLLPLPREAARAAVAAVSGSAAEAK  
ECDRPGNGRCNPGTGCQVCPGWGVEQCHQGRFLTGSSGFVDPGNTKYKFK  
CTWLEQPNKIMRLRFNFHAFETCSWDHLYVDGSIYAPLIAAFSLGVIPERDNET  
APEVTVSGDIAHFFSDAAYNLTFNFDMPNCSGRGECSSNSSSTVECEC  
SENWKBSCDIPHTDNCGPHRGINASDTRGSCFPHWQPGCSIPVPANOSFWTR  
EYSDALPRASHKAVNGNIMVVGGMFNHSDYSMLVAYDLTSREWLPLNHSVNSV  
VRYGHSALHDKIYMYGKIDGNTGNVNELEFVHNESSWLLTPKAKDOXAVVGH  
SAHIVTUSGRVNLVIFGHICPLIYGYISVVOEVDLEKNTWSILHTQGLVQGGYGH  
VYDHTKALYHVGYSKAFKANKYRLADDLRYDVTOMTILKDSRFFRYLHTAVLS  
GTMVLFEGNTHNTDSHSHGAKCFSSDFMAYDIACDRSVLPRLHHDVNRFGHSALV  
HNSTWYVGGFNSLLSLVLFVTSQCDAHRSEAAVCAAGPIRCLWDVQSRCTSWEL  
LAYEQAKLSECFKSRKTLDRCHDOHTCYCTANTINDCHWCHDVPVNHSCTEG  
QISTAKENYDNALSKRNHAFSLASTSQKVELVAKOLRLMQSSQMSKLTLPWV  
GNGLKITTAKENYDNALSKRNHAFSLASTSQKVELVAKOLRLMQSSQMSKLTLPW  
TPVGLKINVTWCHMEDMFFNTSLUOWPSPSDAGFCGLISESTEGSLKATCIN  
PLNGSVGERPANHSKQRTPCALRTACGECTSSSCSNKQCVDSNAYVAFPPGQC  
FQCMWYTSPPENSCSYCTCSHGLEQGGWCTDPSNTGKICEGYSKYPVMPQSA  
SAGNVPQPLNSMCLDSRYNWSFHCPCQKSHKINQSCVCEKEDLTGKHG  
KHCETCISGFVGDPTNGGCKQPKCNHSHASLNTGKFCFTTKGVKGDCEOLCKVE  
NRYGNPLKCTCYVTLIDYQFTFSLSQGDRVYTAINFVATPDEQNRDPMFINAK  
NFNLTWASFPAGTGTGEVFPVSKTNIKEYKDSFSEKEDPFRNHPNITFPVYVNS  
FTWPIKIQIAPFSQHSNFDVJQVFFVFFSCLLSLLVAAVWKIKOSCNASRRERQ  
REMQASRPFASVNVVALEDEPDLIGSIKTVKPIALEPCFNGKAAVLSVFLPRGL  
GGTTPPQSGSLAVASALVDISQMPVIYVKEKSGAVNRKXQPPAQPGTCTI"

ORIGIN

Alignment Scores:

Pred. No.:	3,21e-187	Length:	8739
Score:	206.00	Matches:	568
Percent Similarity:	95.62%	Conservative:	0
Best Local Similarity:	95.62%	Mismatches:	13
Query Match:	14.42%	Indels:	26
DB:	10	Gaps:	0

US-09-787-097-12 (1-1429) x AB038387 (1-8739)

QY	849	SerLysLeuThrLeuThrProTrpValGlyLeuArgLysLeuValSerTyrTrpCys	868
DB	2614	TCCAGCTCAGTCTGACTCCATGGGTGGTCTTCGAGAGATCAATGTATCTTACTGGTGC	2673
QY	869	TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer	888
DB	2674	TGGGAGGATATGCTCCATTCAAAATAGTTTGTGCAAGTGCATCTGAGCCAGT	2733
QY	889	AspAlaGlyPheCysGlyLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr	908
DB	2734	GATCGGGCTTCTGTGGGATCTGTGAGAGCTTACTTCGGGACTGAAGCTCAACC	2793
QY	909	CysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerAlaLysGln	928
DB	2794	TGCATCAACCACTCAATGGAAGTGTGTGAAAGCCCTGCAACCAAGTCCCAAGCAG	2853
QY	929	CysArgThrProCysAlaLeuArgThrAlaCysGlyAsp-CysThrSerGly-SerSerG	948
DB	2854	TGCCGGACACCCCTGTGCCCTGGGACAGCATGTGGTGA-GTGACACAGCAG-CAGCTCCG	2911
QY	948	IuCysMetTrpCysSerAsnMetLysGlnCysValAsnSerAsnAlaTrpValAlaSerP	968
DB	2912	AATGATGTGGTGGAGCAATGAGCAGTGGTGTAGCTTCAATGCTTACGTAGCTCTCT	2971
QY	968	heProPheGlyGlnCysMetGluTyrTrpThrMetSerThr-CysProProGluAsnCys	987
DB	2972	TTCTTTTCGGCAGTGTATGGAGTGTATACAAATGAGCAG-CTGCCACCTGAAATATGC	3030
QY	988	SerGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyTyrCysThrAsp	1007
DB	3031	CTGTGCTACTGTACTGTGAGCATTGCTTGGAGCAGCAGGCTGGTGGTGTACTGAT	3090
QY	1008	ProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMet	1027
DB	3091	CTTAGCAATCTGGGAAAGGCAATGATTTAGGGAGCTATAAGGACCCCTGGAATG	3150
QY	1028	ProSerGlnAlaPro--ThrGlyAsnPhe-TyrProGlnProLeuLeuAsnSerSerMet	1046
DB	3151	CCTTCACTG-CCTCTACAGGAAATGT-GTACCCACAGCCCTTCTGAACCTCCAGCATG	3207
QY	1047	CysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsn	1066
DB	3208	TGTCTAGAGACAGCAGATACAACTGCTCTTTTATTCAGTCTCCAGCTTGGCAGTGAAT	3267
QY	1067	GlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn-LeuThrThrG	1086
DB	3268	GGACACAGCAAGTGCATCAACAGACATCTGTGAGAAGTGTGAGA-CCTGACACAGG	3326
QY	1086	LysHisCysGluThrCysIleSerClyPheTyrGlyAspProThrAsnGlyLysCys	1106
DB	3327	CAAGCACTCGAGACCTGATATCTGGCTTCTATGTCACCGACTAATGAGGCAATG	3386
QY	1106	sGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPh	1126
DB	3387	TCAGCCATGCAAGTGCAACGGGACGACATCACTGTGTCAACCAACACTGGCAAGTCTT	3446
QY	1126	eCysThrThrLysGlyValLysGlyAsp-GluCysGlnLeuCysGluValGluAsnArgT	1146
DB	3447	CTGTACCAACAGGGCGTCAAGGAGA-GGAGTGGCAGCTATGTAGGTAGAAAATCGAT	3505
QY	1146	ValGlnGlyAsnProLeuArg-GlyThrCysTyrThrLeuLeuIleAspTyrGlnPhe	1165
DB	3506	ACCAGGAAACCTCTCAA-AGGACATGTATTACTTACCTCTCTATTGACTATCAGTTC	3564

QY	1166	ThrPheSerLeuSerGlnGluAspAspArgTyrThrAlaIleAsnPheValAlaThr	1185
DB	3565	ACTTTTAGCTGTCCAGGAGAGCAGCCGCTATTACACAGCCATCAATTTTGGCTACT	3624
QY	1186	ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu	1205
DB	3625	CCTGATGAACAAACAGGATTTGGACATGTTTCATCAATGCTCCCAAAATCTTCAACCTC	3684
QY	1206	AsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGlnAla-GlyGluGluMet-Pr	1224
DB	3685	AACATACCTTGGGCCAC-CAGTTTCCC-AGCTGGNACCCAGAC-TGAGAGAGAGGT-GCC	3740
QY	1224	oValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGlnLysPheAs	1244
DB	3741	TGTTGTTTCAAAACCAACATCAAGGAATCAAAAGATAGCTTCTCTAATGAGAAATTGA	3800
QY	1244	pPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpProIl	1264
DB	3801	TTTTTCGCAACCATCCAAACATTTCTTTTGTGTATGTCAATTAATTTCACTTGGCCCAT	3860
QY	1264	eLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheVa	1284
DB	3861	CAAAATTCAAATTCCTTCTCCAGCAGCAGCAATTTTATGGACCTGGTGCAGTCTTCTCGT	3920
QY	1284	lThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrpLysIleLys	1304
DB	3921	GACTTTCTTCACTGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	3980
QY	1304	sGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAl	1324
DB	3981	GCAGAGCTGTGGGCATCCAGGAGGAGAGCAACTTCTTCGGGAGATGCAACAGATGGC	4040
QY	1324	aSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluProProAspLe	1344
DB	4041	CAGCGCCCTTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4100
QY	1344	uIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyAs	1364
DB	4101	CATTGGGGAGATATAAGACTGTTCTTAAGCCCATTTGCCCTGGAGCCCTGCTTTGTAA	4160
QY	1364	nLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyIleProPr	1384
DB	4161	CAAAAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	4220
QY	1384	oProGlyGlnSerClyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetPr	1404
DB	4221	TCCTGGTCAGTCAGTCTTGTGGCCAGTGGCCCTGGTGGACATTTCTCAGCAGATGCC	4280
QY	1404	oIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGl	1424
DB	4281	AATAGTGTACAGGAGAGATCAGGAGCTGTAGGAACCGGAGAGCAGCAGCCCTTGACA	4340
QY	1424	nProGlyThrCysIle 1429	
DB	4341	GCCTGGAACCTGCATT 4356	
RESULT 13			
LOCUS	AR164807	8827 bp	DNA linear PAT 17-OCT-2001
DEFINITION	Sequence 1 from patent US 6274339.		
ACCESSION	AR164807		
VERSION	AR164807.1	GI:16238010	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 8827)		
AUTHORS	Moore,K. and Nagle,D.Lynn.		
TITLE	Methods and compositions for the diagnosis and treatment of body		
	weight disorders, including obesity		
JOURNAL	Patent: US 6274339-A 1 14-AUG-2001;		
FEATURES	Location/Qualifiers		
	1..8827		
	source		

/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

## Alignment Scores:

Pred. No.: 3,23e-187 Length: 8827  
Score: 206.00 Matches: 576  
Percent Similarity: 96.00% Conservative: 0  
Best Local Similarity: 96.00% Mismatches: 12  
Query Match: 14.42% Indels: 24  
DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x AR164807 (1-8827)

QY 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861  
DB 2790 ATGCAATCATCTCAAGATGTCCAGCTCACTGCTCACTCCATGGGTGGTCTTCGGAAG 2849  
QY 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881  
DB 2850 ATCAATGTGTCTTACTGGTCTGGGAGGATATGTCTCCATTCAAAATAGTTTGTCTGCAG 2909  
QY 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyLeuLeuSerGluProSerThr 901  
DB 2910 TGGATGCCATCTGAGCCAGTGTGCTGGCTCTGTGGGATCTTGTACAGAGCTAGTACT 2969  
QY 902 ArgGlyLeuLysAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
DB 2970 CGGGGATTAAGGCTGCACTGATCATCACTCTCAATGGCAGCTCTGTGAAGGCT 3029  
QY 922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941  
DB 3030 GCAAAACACAGTCCAGCAGTGTCCGGACACCATGTGCCCTGGGACAGCGTGTGGCGA- 3088  
QY 942 -CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS 961  
DB 3089 GTGCACCTAGCAG-CAGCTCGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3147  
QY 961 erAsnAlaTyrValAlaSerPheProPheCysMetGlnCysMetGluTrpTyrThrMetSerT 981  
DB 3148 CCAATGCTACGTGGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3207  
QY 981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000  
DB 3208 G-CTGCCACCTCAAAATGTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 3266  
QY 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020  
DB 3267 GGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3326  
QY 1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039  
DB 3327 TATTAAGGACCTGTGAAGATGCCGTACAGGCTC-TG-CAGGAATGT-GTATCCACAG 3383  
QY 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheLeuHis 1059  
DB 3384 CCGCTTCTGAACCTCCAGCATGTCTAGAGGACAGCAGATACAACCTGGTCTTCTATTCTAC 3443  
QY 1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079  
DB 3444 TGTCCAGCTTGCAGTGCACGACACAGCAATGTGATCAACAGAGTATCTGTGAGAG 3503  
QY 1080 CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099  
DB 3504 TGTGAGGA-CCTGACACCGGGCAAGCACTGCGAGACTGCGATATCTTGTGGTCTTATGTGTA 3562  
QY 1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119  
DB 3563 CCGGACTAATGGAGGAAATGTGAGCCATGCAATGTGGCAGCGCATCTGTGCA 3622  
QY 1119 nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139  
DB 3623 CACCAACACCGGCAAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3682

QY 1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159  
DB 3683 ATGTGAGGTAGAAATCGATACCAAGGAACCTCTCAA-AGGAACATGCTACTATACCC 3741  
QY 1159 euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspTyrTyrThrA 1179  
DB 3742 TTCTCATGTGATCATCAGTTACCTTTAGCCTGTCCAGGAAGACGACCGCTACTACACG 3801  
QY 1179 laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199  
DB 3802 CCATCAACTTGTGCTACTCTCTGATGAACAAACAGGGATTTGGACATGTTTCATCAATG 3861  
QY 1199 laSerLysAsnPheAsnLeuAsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGI 1218  
DB 3862 CTTCAAAATCTTCAACCTCAACATCACTGGGCCAC-CAGCTTCCC-AGCCGGAACCCA 3919  
QY 1218 nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237  
DB 3920 CAC-TGGAGAGAGGT-GCCTGTGTGTTTCAAAACCAACATCAAGGATACAAAGATAGC 3977  
QY 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257  
DB 3978 TTCTCTAATGAGAAATTTGATTTTTCGCAACCATCTCAAAACATCACTTTCTTTTGTATGC 4037  
QY 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277  
DB 4038 AGTAATTTCACTTGGCCCATCAAAATTCAGATGCTTCTCCAGCACCACTTTCATG 4097  
QY 1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAla 1297  
DB 4098 GACCTGGTACAGTTCTTCTGTGACCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4157  
QY 1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeu 1317  
DB 4158 GCAGTGTGTGGAAGATCAAGCAGAGCTGTGGCATCCAGGGGAGAGAGCACTTCTT 4217  
QY 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337  
DB 4218 CGGAGATGCACAGATGGCCGCGCTTCTTCTGTAACCGTTGCTTGGAAACA 4277  
QY 1338 AspGluGluProAspLeuIleGlySerIleLysThrValProLysProIleAla 1357  
DB 4278 CATGAGGAGCTCTGTGATCTTATTTGGGGGAGTATAAGACTGTTCCTCAACACCATTCGA 4337  
QY 1358 LeuGluProCysPheGlyAsnLysAlaValLeuSerValPheValArgLeuProArg 1377  
DB 4338 CTGGAGCGGTGTTTGGCAACAAGCCCTGTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 4397  
QY 1378 GlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397  
DB 4398 GGCTGGTGTGTCATCCCTCTCTCTGGGCACTCAGGTCTTCTGTGTGTGTGTGTGTGTGTGTGTGT 4457  
QY 1398 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArg 1417  
DB 4458 GACATTTCTCAGCAGATGCGGATAGTGTACAGGAGAGTACAGAGCCGTGAGAACCCGG 4517  
QY 1418 LysGlnGlnProAlaGlnProGlyThrCysIle 1429  
DB 4518 AAGCAGCAGCCCTCTGCACAGCTGGGACCTGCATC 4553  
RESULT 14  
AB093245 5683 bp mRNA linear ROD 07-FEB-2003  
LOCUS Mus musculus mRNA for mKIAA0548 protein.  
DEFINITION AB093245  
ACCESSION AB093245.1 GI:26006172  
VERSION FLI CDNA.  
KEYWORDS Mus musculus (house mouse)  
SOURCE Mus musculus  
ORGANISM Mus musculus  
REFERENCE 1  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.

**AUTHORS** Okazaki, N., Kikuno, R., Ohara, R., Inamoto, S., Hara, Y., Nagase, T., Ohara, O. and Koga, H.

**TITLE** Prediction of the coding sequences of mouse homologues of KIAA gene: I. The complete nucleotide sequences of 100 mouse KIAA-homologous cDNAs identified by screening of terminal sequences of cDNA clones randomly sampled from size-fractionated libraries

**JOURNAL** Unpublished

```

1. 5693
/gene="mKIAA0548"
<1. 3181
/gene="mKIAA0548"
/note="CDS is predicted by in silico analysis. Start codon
is not identified."
/codon_start=2
/evidence=not experimental
/product="mKIAA0548 protein"
/protein_id="BAC41429.1"
/db_xref="GI: 26006173"
/translation="AYDLTSEWVPLNHSVNSVVVRYGSLALEHKDKIYMYGCKIDST
GNVNRVRFHINESVWLLTPKAKQAVVGHSAHIVTLASGRVVMVLVIYFGHCPLYIG
YISVQVEYDEKNTWSLHTQAGLVQGGYSHSYVDRTKALYVGGYKAFSAANKYRL
ADLYRVYDVDTQMTWLTGFFLYLTAIVTSGVTLVFGCNGTHNTDSMGHAKCFSS
DFMAYDIACDRWSLVLPRELHDDNVRFGHSAVLNSTMYVFGFNSLLLSDLVLFVLE
QCDHRSBAACVAAGPVGIRCLMDTQSSRCTSWELATEBOAEKLSKCFSKRTLDHRC
DOHTDVCYSTANDCHNCVDPVSHSTEGGSIATKYESCPKSCPMFTYCNKNTSC
RSCALDQCNREPOEQTALPENICGNHMLVGNLSCLITAKENYDNAKLSCRNH
AFLASLTQKQVEFKLQRLQMSQSGMLTTPVGLRAKINYSVMEDMSPFPFNS
LLOWMPSPSPDAGFCGLISBSTGLKAAATCINPLGSCVCEPANHSAKOCRTPCALR
TAGCEBTSSSECCWCSNKQCVDSNAYVSPFPQGCWENYTWSSCPENCSCYCTCS
HCLBPQGGCWTDSNTGKGCISYSGYKGPVMSQASAGNVPQPLNLSMCLDSCR
YNWSIHPCAQCNHGSCINQSCVEKEDLTGKHCTCISGIDPTNGKQCKQPC
CNGHASLNTWGTGCTCTKGAGGECQLCEVENYQCNPLKGTCTYTLIDYQTFPS
LSQEDDRTYTAINFVATEQDNRLDMFINASKFNELNITWATSPAGTGTGSEVPV
KFTTCLEKYDQSFSEKEDFRHNPHITFPVYVSNFTWPIKIQIAFSOHNVALETDERPV
SFFCSKTELLILVAAWKIKSCWASRRREQLLREMQMASRFPASVNVVALETDERPVL
DLIGSIIKTPKTLAETPCFGNKAALVSFVRLPRLGJGIPPPQGSGLAVASALVDVIS
QCMPTVYKESGAVNRKQQPPAQGTGTC"

```

## ORIGIN

Alignment Scores:	
Pred. No.:	Length: 5693
Score:	Matches: 574
Percent Similarity:	Conservative: 0
Best Local Similarity:	Mismatches: 14
Query Match:	Indels: 28
DB:	Gaps: 0
	10
	1.25e-180

US-09-787-097-12 (1-1429) x AB093245 (1-5683)

Qy 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861  
Db 1415 ATGCAATCATCTCAAAAGTATGTCCAAAGTCCTCTGACTCCATGGGTGGTCTTCGGAAG 1474

Db	2544	AGAC-TGACAGAGGT-GCTGTGTTTCAAAACAC-CATCAAGGAATACAAAGATA	2600
Qy	1237	erPheSerAnGluLysPheAspPheArgAnHisProAnleThrPhePheValTyrV	1257
Db	2601	GCTTCTCTAATGAGAAATTTGATTTGGCAACCATCAACATCACTTCTTTGTTATG	2660
Qy	1257	alSerAnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAnPheM	1277
Db	2661	TCAGTAATTCACCTGGCCCATCAAAATTCAGATTCCTTCTCCAGCAGCACTTCA	2720
Qy	1277	etAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValA	1297
Db	2721	TGGACCTGGTACAGTCTTCTCGTACCTTCTTCAAGTTGTTCTCTCGCTTCTGGTGG	2780
Qy	1297	laAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGluGlnLeuL	1317
Db	2781	CTCAGTGTCTGAGATCAACAGAGCTGTGGCATCCAGCCGAGAGAGCACTTC	2840
Qy	1317	euArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluT	1337
Db	2841	TTCCGGAGATGCAACAGATGGCCGAGCCCTTCTGCTTCTGTAACGTTGCTTGGAA	2900
Qy	1337	hrAspGluGluProProAspLeuIleGlySerIleLysThrValProLysProIleA	1357
Db	2901	CAGATGAAGAACCTCTCTGATCTCATTTGGGGAGATATAAGACCGTCTTAAGCCCATG	2960
Qy	1357	laLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProA	1377
Db	2961	CCCTGGAGCCCTGTTTGGTAACAGCCGAGTCTCTCTGATTCGTGAGCTCCCTC	3020
Qy	1377	rgGlyLeuGlyGlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuV	1397
Db	3021	GAGGACTGGGAGGAATCCCTCTCTCTGTCAGTCAAGTCTCGTGTGGCCAGTCCCTGG	3080
Qy	1397	alAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnA	1417
Db	3081	TGGACATTTCTAGCAGATGCCATAGTGTACAGGAGAGTCAAGAGCTGTAGAAACC	3140
Qy	1417	rgLysGlnGlnProProAlaGlnProGlyThrCysIle	1429
Db	3141	GGAAGACAGACCGCCCTGTACAGCCCTGCAACCTGCATT	3178
RESULT 15			
LOCUS	AF116897	6083 bp mRNA linear	ROD 20-MAR-1999
DEFINITION	Mus musculus mahogany protein mRNA, complete cds.		
ACCESSION	AF116897		
VERSION	AF116897.1	GI:454560	
KEYWORDS			
SOURCE	Mus musculus (house mouse)		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Nagle, D.L., McGrail, S.H., Vitale, J., Woolf, E.A., Dussault, B.J. Jr., DiRocco, L., Holmgren, L., Montagnon, J., Bork, P., Huszar, D., Fairchild-Huntress, V., Ge, P., Keilty, J., Ebeling, C., Baldini, L., Gilchrist, J., Burr, P., Carlson, G.A. and Moore, K.J.		
TITLE	The mahogany protein is a receptor involved in suppression of obesity		
JOURNAL	Nature 398 (6723), 148-152 (1999)		
MEDLINE	99184159		
PUBMED	10086355		
REFERENCE			
AUTHORS	Nagle, D.L., McGrail, S.H., Vitale, J., Woolf, E.A., Dussault, B.J. Jr., DiRocco, L., Holmgren, L., Montagnon, J., Bork, P., Huszar, D., Fairchild-Huntress, V., Ge, P., Keilty, J., Ebeling, C., Baldini, L., Gilchrist, J., Burr, P., Carlson, G.A. and Moore, K.J.		
TITLE	Direct Submission		
JOURNAL	Submitted (28-DEC-1998) Sequence Analysis, Millennium Pharmaceuticals, 640 Memorial Drive, Cambridge, MA 02139, USA		
FEATURES	Location/Qualifiers		

source	1. 6083		
	/organism="Mus musculus"		
	/mol_type="mRNA"		
	/db_xref="taxon:10090"		
CDS	70..4488		
	/note="MG"		
	/codon_start=1		
	/product="mahogany protein"		
	/protein_id="AAD20947.1"		
	/db_xref="GI:4454561"		
	/translation="MVAFAAAATEARLGRSTTTTAAPAGKGRGHRPCTATCAWRPG FRALGCRPVISRALPPPLPLLLPLLLPRLPRAEAAAVAAVSGSAAAEKGRDR PFGGCRNPGTCCQCTGPTGWEQCHQGRFRLTSSGFTVDTGPNYKTKTKTWL IEQPNRIMKLRNFHATECSWDHLVYDGLIYAPLIAFSLIYPERDGNETAPEV TVTSGYALLHFFDAAYNLTFNITVDMPCNNSGRGCKSSNSVASCSESNW KGSCDIPHTCNDGCPHRRGICNASPTRGSCFPHWQPGCSI PVANOSFWTRYS DLKLPRAHKAIVNGIMVWVGGYMHNSDYSMVLADLTREMLPLNHSVSVVRY GHSIALHKKIYMGKIDSTGNVTHLRVFLHNSWLLTTPKADQYAVVGHSAHI VTLASGRVMLVIFGHCPLYISVQEXYDLEKNTWSILHTQALVGGTGHSSVTD RTKALYVHGKYKAFSAKSKRLADLYRVDYDVTQMTILKDSRFRYLHTAVIVSGTMD VFGGNTDNTSMGSAKCFSSDFMAYDIACRMSVLPRLPRLHVDVRFHSAVLYNST MYVFGGNSLLSLVFTSEQDAHRSAAACVAAGPIRCLWDTQSSRCTSWELATE EQAEKLSKCFSKRTLDHSCDOHTDCYSTANTNDCHWCHDNCVPCNRSCTBGQSI AKYESCPKDPMYCKNKTSCRCALDONCEPRNOECIALPENICNGWHLVNGSC LKITTAKENDYAKLSRNHNFLASLTQKKVEFVLKQLRLNQSQSSSKLTLTPWY GLRKINYSYWCWEDMSPTNLSLWMPSEPSDAGFCILSEPTSGLKAAATCINPLNG SVCEPANHSAKQCTPTCALRTAGCECTSSSECMCNKOCVSNAYVACVFPQOC MEWTVSSCPENCSCYCTCHLEQPGCWCTDPSNTGKGIKIGSYKGPVMPQSC SAGNVYPOPLINSMLCLESRYNWSFTHCPACONGHSKINOSICEKEDILTGHKC ETICSGPYGPTNGKCPCKNGHSLQNTNTGKCTCTTKGVKGBECOLCEVENEYQ GNPLKCTCYTLLIDYQFSLSQEDDRYTAINFVATPDNEONEDLDMENASKNENL NITWASFPAGTQGEVSVKINIKYKOSFNSNEKPFPRHNPNTIFFVYSNFTWP IKIQAFQSHNFMDLVQFVTFPFCFLSLLVAVAVKIKQSCASRRRRLQRLREMQ QMARPAPASVVALETDDEPDLIGSITKVPKPIALEPFCGNKAAVLVSVFLPRL GGTPPQSRSGCQCGPHFSADANSVQEVRSCKPEAAAACTAWNLHLILUGQFPA LTEGWSTKLSGTGERAGRKQAWGRRLLGTLEASNSCA"		
ORIGIN			
Alignment Scores:			
Pred. No.:	7,86e-175	Length:	6083
Score:	193.00	Matches:	576
Percent Similarity:	95.84%	Conservative:	0
Best Local Similarity:	95.84%	Mismatches:	12
Query Match:	13.51%	Indels:	25
DB:	10	Gaps:	0
US-09-787-097-12 (1-1429) x AF116897 (1-6083)			
Qy	842	MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys	861
Db	2590	ATGCAATCATCTCAAGTATGTCAAGCTCACCTGACTCCATGGTGTGCTTCGGAAG	2649
Qy	862	IleAsnValSerTyrTyrCysTrpTyrGluAspMetSerProPheThrAsnSerLeuGln	881
Db	2650	ATCAATGTGCTTACTGGTCTGGAGGATATGTCCATTCACAAATAGTTTGTGTCAG	2709
Qy	882	TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr	901
Db	2710	TGGATGTCATCTGAGCCAGTGAIGTGGTCTTGTGGGATCTGTGACAGCCCTAGTACT	2769
Qy	902	ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro	921
Db	2770	CGGGGATTAAGGCTGCAACCTGATCAACCCCTCTCAATGGCAGCGCTCTGTGAAAGCCCT	2829
Qy	922	AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp	941
Db	2830	GCAACCCACAGTGCCACAGAGTCCCGACACCATGTGCCCTGGCAGACGGTGTGGCA	2888
Qy	942	-CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp	961
Db	2889	GTGCACCTAGCAG-CAGCTCGAGTGCATGTGTGTGTCAGTAAACATGAAGCAGTGTGTGACT	2947
Qy	961	erAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrThrMetSerT	981

Db 4018 CGGAGNATGCAACAGATGGCCAGCCGCCCTTGTCTCTTAAACGTTGCCTTGGAACA 407  
 Qy 1338 AspGlucludProProAspLeuLeuGlyGlySerIleLysThrValProLysProIleAla 1357  
 Db 4078 GATCAAGAACCTCTCTGATCTCATTTGGGGGAAGTATAAGACCGTTCTTAAGCCCATGGC 4137  
 Qy 1358 LeuclnProCysPheGlyAsnLysAlaValLeuSerValPheValArgLeuProArg 1377  
 Db 4138 CTGAGGCCCTTGCTTTGTATAAAGCCGCACTCTCTCTGTATTCTGTGAGGCTCCCTCGA 4197  
 Qy 1378 GlyLeuGlyGlyLeuProProGlyGlnSer-GlyLeuAlaValAlaSerAlaLeuVa 1397  
 Db 4198 GGACTGGAGGAATCCCTCTCTCTGTCAGTCAAGGTCCTCGCTGTGGCCAGTGCCTGGT 4257  
 Qy 1397 AspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnAr 1417  
 Db 4258 GGACATTTCTCAGCAGATGCCAATAGTGTACAGGAGAGTCAAGGACTGTAAAGAAACCG 4317  
 Qy 1417 gLysGlnGlnProProAlaGlnProGlyThrCysIle 1429  
 Db 4318 GAAGCAGACGCCCTGACAGCCTGGAACTGCATT 4354  
  
 RESULT 16  
 AX778545  
 LOCUS  
 DEFINITION Sequence 702 from Patent WO03039443.  
 ACCESSION AX778545  
 VERSION AX778545.1 GI:326955539  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1  
 REFERENCES  
 AUTHORS Haferlach, T., Schoch, C., Kern, W., Kohlmann, A., Schnittger, S.,  
 Dugas, M., Ellis, R., Brors, B. and Mergenthaler, S.  
 Novel genetic markers for leukemias  
 Patent: WO 03039443-A 702 15-MAY-2003;  
 Deutsches Krebsforschungszentrum (DKFZ)  
 Ludwig-Maximilians-Universität München (LMU)  
 PD Dr. Dr. (DE) ; Schoch, Claudia (DE) ; Kern, Wolfgang (DE)  
 FEATURES  
 source  
 1..543  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 4,05e-164 Length: 543  
 Score: 181.00 Matches: 181  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 12.67% Indels: 0  
 DB: 6 Gaps: 0  
  
 US-09-787-097-12 (1-1429) x AX778545 (1-543)  
 Qy 1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103  
 Db 1 ACCACAGCGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAACATGGA 60  
 Qy 1104 GlyLysCysGlnProCysLysCysAsnGlyHisIleSerLeuCysAsnThrAsnThrGly 1123  
 Db 61 GGGAAATGTTCAGCCATCGAATGGGACCGCTCTCTGTGTCACCAACCAACACGGGC 120  
 Qy 1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143  
 Db 121 AAGTGCCTTCTGCACCAACAGGGCGTCAAGGGGACGAGTGCCAGCTATGTGAGGTAA 180  
 Qy 1144 AsnArgTyrGlnGlnAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyr 1163  
 Db 181 AATCGATACCAAGAAACCCCTCTCAGAGAACATGTTATTATATCTTCTTAATGACTAT 240



QY 1164 GlnPheThrPheSerLeuSerGlnGluAspArgTyrThrAlaIleAsnPheVal 1183  
 |||||  
 Db 241 CAGTTCACTTTAGTCTATCCGAGGAGATCGCTATTACACAGCTATCAATTTGTG 300  
 |||||  
 QY 1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerIlyAsnPhe 1203  
 |||||  
 Db 301 GCTACTCCTGACCAACAAACAGGAGATTGGACATGTTCAATCAATGCTCCCAAGAAATTC 360  
 |||||  
 QY 1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluMet 1223  
 |||||  
 Db 361 AACCTCAACATCACTCGGGCTGCCAGTTCTTCAGCTGGAAACCCAGGCTGGAGAGATG 420  
 |||||  
 QY 1224 ProValValSerIlyThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243  
 |||||  
 Db 421 CTTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTCTCTAATCAGAAATTT 480  
 |||||  
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263  
 |||||  
 Db 481 GATTTTGGCAACCAACCAATATCACTTCTTTGTTTATGTGATTAATTCACCTGGCC 540  
 |||||  
 QY 1264 Ile 1264  
 |||||  
 Db 541 ATC 543

# RESULT 17 BX004762\_0/c WPCOMMENT

Sequence split into 5 fragments LOCUS BX004762 Accession BX004762

Fragment Name	Begin	End
BX004762_0	1	110000
BX004762_1	100001	210000
BX004762_2	200001	310000
BX004762_3	300001	410000
BX004762_4	400001	442165

LOCUS BX004762 442165 bp DNA linear HTG 12-APR-2003  
 DEFINITION Mus musculus chromosome X clone RP23-146C4, WORKING DRAFT SEQUENCE,  
 4 unordered pieces.

ACCESSION BX004762  
 VERSION BX004762.6 GI:29824756  
 KEYWORDS HTG; MUSCULUS; HTGS\_ACTIVEFIN; HTGS\_DRAFT; HTGS\_FULLTOP.  
 SOURCE Mus musculus (house mouse)  
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 442165)

## REFERENCE AUTHORS TITLE JOURNAL

Direct Submission  
 Submitted (11-APR-2003) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Apr 13, 2003 this sequence version replaced gi:29125168.  
 Sequence from the Mouse Genome Sequencing Consortium whole genome  
 shotgun may have been used to confirm this sequence. Sequence data  
 from the whole genome shotgun alone has only been used where it has  
 a phred quality of at least 30.

----- Genome Center  
 Center: Wellcome Trust Sanger Institute

Center code: SC  
 Web site: <http://www.sanger.ac.uk>

Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)  
 ----- Project Information

Center project name: bM146C4  
 ----- Summary Statistics

Assembly program: XGAP4; version 4.5  
 Chemistry: Dye-terminator; 100% of reads

Consensus quality: 229033 bases at least Q40  
 Consensus quality: 229393 bases at least Q30

Consensus quality: 229535 bases at least Q20  
 Insert size: 441865; sum-of-contigs

Insert size: 188914; 6.6% error; agarose-fp  
 Quality coverage: 3.34x in Q20 bases; sum-of-contigs Quality

coverage: 7.96x in Q20 bases; agarose-fp

----- NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

\* 1 48919: contig of 48919 bp in length  
 \* 48920 49013: gap of 100 bp  
 \* 49020 130781: contig of 81762 bp in length  
 \* 130782 130881: gap of 100 bp  
 \* 130882 229964: contig of 99083 bp in length  
 \* 229965 230064: gap of 100 bp  
 \* 230065 442165: contig of 212101 bp in length.

## FEATURES Location/Qualifiers

source  
 1..442165  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10090"  
 /chromosome="X"  
 /clone="RP23-146C4"  
 /clone\_lib="RPC1-23"  
 1..48919  
 /note="assembly\_fragment:00147"  
 fragment\_chain:1  
 49020..130781  
 /note="assembly\_fragment:00500"  
 fragment\_chain:1  
 130882..229964  
 /note="assembly\_fragment:00870"  
 fragment\_chain:1  
 230065..442165  
 /note="assembly\_fragment:02147"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3.59e-113 Length: 110000  
 Score: 130.00 Matches: 130  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 9.10% Indels: 0  
 DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x BX004762\_0 (1-110000)

QY 1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278  
 |||||  
 Db 101355 AATTTCACCTTGGCCCATCAAAATTCAGATTGCCTTCTCCAGCACACAGCACTTCATGGAC 101296  
 |||||  
 QY 1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaLeu 1298  
 |||||  
 Db 101295 CTGTACAGATCTTCGAGACTTTCTTCAGTTGTTTCTCTCGCTGCTTCGTGGGTGCA 101236  
 |||||  
 QY 1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArg 1318  
 |||||  
 Db 101235 GTGGTCTGGAAGATCAAGACAGAGCTGTGGCATCCAGCGGAGAGAGCAACTTCTTCGG 101176  
 |||||  
 QY 1319 GluMetGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp 1338  
 |||||  
 Db 101175 GAGATGCAACAGATGGCCGCCCTTTGCTTCTGTAAACCTTGCCTTGGAAACAGAT 101116  
 |||||  
 QY 1339 GluGluProAspLeuIleGlySerIleLysThrValProLysProIleAlaLeu 1358  
 |||||  
 Db 101115 GAGAACCTCTCTGATCTCATTTGGGGAAGTATAAGACCGTTCCTTAAGCCATTGCCCTG 101056  
 |||||  
 QY 1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly 1378  
 |||||  
 Db 101055 GAGCCCTGCTTTGGTAAACAAAGCCGAGTCTCTCTGTATTCTGTAGGCTCCCTCGAGGA 100996  
 |||||  
 QY 1379 LeuGlyGlyIleProProGlyGlnSer 1388  
 |||||



Db 100995 CTGGAGGAATCCCTCCTCTCGTCAGTCA 100966

RESULT 18  
BX004762\_1/c  
WPCOMMENT

Sequence split into 5 fragments LOCUS BX004762 Accession BX004762

Fragment Name	Begin	End
BX004762_0	1	110000
BX004762_1	100001	210000
BX004762_2	200001	310000
BX004762_3	300001	410000
BX004762_4	400001	442165

Continuation (2 of 5) of BX004762 from base 100001 (BX004762 Mus musculus chromosome X)

Alignment Scores:

Pred. No.:	Score:	Length:
3.59e-113	130.00	110000
Percent Similarity:	100.00%	Matches: 130
Best Local Similarity:	100.00%	Conservative: 0
Query Match:	9.10%	Mismatches: 0
DB:	2	Indels: 0
		Gaps: 0

US-09-787-097-12 (1-1429) x BX004762\_1 (1-110000)

Qy	AsnPheThrTrpProIleLysGlnIleAlaPheSerGlnHisSerAsnPheMetAsp	1278
Db	1355 AATTTCACCTTGCGCCCATCAAATAATTCAGATTGCCCTTCCCCAGCACACAACCTTCATGCAC	1296
Qy	1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAala	1298
Db	1295 CTGGTACAGTTCTTCGTGACTTTCTTCAGTTGTTCTCTCGCTGCTTCGTGGTGGTGCA	1236
Qy	1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArg	1318
Db	1235 GTGGTCTGGAAGATCAGACAGAGCTGTGGGCATCCAGCGGAGAGCAACTTCTTCGG	1176
Qy	1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp	1338
Db	1175 GAGATGCAACAGATGGCGCGGCCCTTGTCTTCTCTAAACGTTGCTTGGAAACAGAT	1116
Qy	1339 GluGluProProaspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeu	1358
Db	1115 GAAGAACCCTCTGATCTCANTGGGGGAAGATATAAGACCGTTCCTTAAGCCCATTCGCCCTG	1056
Qy	1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly	1378
Db	1055 GAGCCCTGCTTTGGTAACAAGCGCAGTCCTCTCTGTTATTCGTGAGGCTCCCTCGAGGA	996
Qy	1379 LeuGlyGlyIleProProGlyGlnSer	1388
Db	995 CTGGAGGAATCCCTCCTCTCGTCAGTCA 966	

RESULT 19  
AL807401  
LOCUS

DEFINITION AL807401 110261 bp DNA linear ROD 31-OCT-2002  
Mouse DNA sequence from clone Rp23-133Cl2 on chromosome X, complete sequence.

ACCESSION AL807401  
VERSION AL807401.11 GI:24474469  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 110261)  
Howden, P.  
Direct Submission  
Submitted (31-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Nov 1, 2002 this sequence version replaced gi:24395101.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute

COMMENT

Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquery@sanger.ac.uk](mailto:humquery@sanger.ac.uk)

-----  
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-133C12 is from the RPCI-23 Mouse PAC library  
constructed by the group of Peter de Jong.  
For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACE3.6.

FEATURES

	source	Location/Qualifiers
	1..110261	
	/organism="Mus musculus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10090"	
	/chromosome="X"	
	/clone="RP23-133C12"	
	/clone_lib="RPCI-23"	

ORIGIN

Alignment Scores:			
Pred. No.:	3,59e-113	Length:	110261
Score:	130.00	Matches:	130
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	9.10%	Indels:	0
DB:	10	Gaps:	0

US-09-787-097-12 (1-1429) x AL607401 (1-110261)

QY	1259	AsnPheThrTirProIleIysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp	1278
Db	33174	AAATTACATTGGCCCATCAAAATTCAGATTGCCTTCTCCAGACACCAACTTCATGGAC	3323
QY	1279	LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaA	1298
Db	33234	CTGGTACAGTCTTCGTGACTTTCTTCAGTGTGTTCTCTCGTGCCTTCGTGGTGGCTGCA	3329
QY	1299	ValValTirPylsIleIysGlnSerCysTirPAlaSerArgArgGluGlnLeuLeuArg	1318
Db	33294	GTGGTCTGGAGATCAAGACAGCTGTGGCATCCAGCGCGAGACCACTTCTTCGG	3335
QY	1319	GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAsp	1338
Db	33354	GAGATGCACAGATGCCAGCCGCCCTTTGCTTCTCTTAACACGTTCCTGGAAACAGAT	3341
QY	1339	GluGluProProAspIleuIleGlyGlySerIleIysThrValProIysProIleAlaLeu	1358
Db	33414	GAGAACCTCTCTGATCTCATTTGGGGGAAGTATTAAGACCGTTCTTAAGCCCATTTGCCCTG	3347
QY	1359	GluProCysPheGlyAsnIysAlaAlaValLeuSerValPheValArgLeuProArgGly	1378
Db	33474	GAGCCCTGCTTTTGGTAACAAAGCCGAGTCTCTCTGTATTGGTGGGCTCCCTCGAGGA	3353
QY	1379	LeuGlyGlyTleProProProGlyGlnSer	1398
Db	33534	CTGGAGGAATCCCTCTCTCTGTGTCAGTCA	33563

```

RESULT 20
BX005139          335799 bp   DNA      linear   HTG 28-FEB-2003
LOCUS             Mus musculus chromosome X clone RP23-469F12, 3 unordered pieces.
DEFINITION        BX005139
ACCESSION         BX005139
VERSION           BX005139.4 GI:28625284
KEYWORDS          HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE            Mus musculus (house mouse)
ORGANISM          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE         Clark,S.
AUTHORS           Direct Submission
TITLE             Submitted (27-FEB-2003) Wellcome Trust Sanger Institute, Hinxton,
JOURNAL           Cambridgehire, CB10 1SA, UK. E-mail enquiries:
                  humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
                  On Feb 28, 2003 this sequence version replaced gi:28460315.
COMMENT           Sequence from the Mouse Genome Sequencing Consortium whole genome
                  shotgun may have been used to confirm this sequence. Sequence data
                  from the whole genome shotgun alone has only been used where it has
                  a phred quality of at least 30.
                  ----- Genome Center
                  Center: Wellcome Trust Sanger Institute
                  Center code: SC
                  Web site: http://www.sanger.ac.uk
                  Contact: humquery@sanger.ac.uk
                  ----- Project Information
                  Center project name: BM469F12
                  ----- Summary Statistics
                  Assembly program: XGAP4; version 4.5
                  Chemistry: Dye-terminator; 100% of reads
                  Consensus quality: 165573 bases at least Q40
                  Consensus quality: 185718 bases at least Q30
                  Consensus quality: 185822 bases at least Q20
                  Insert size: 335599; sum-of-contigs
                  Insert size: 160496; 4.3% error; agarose-fp
                  Quality coverage: 5.65x in Q20 bases; sum-of-contigs Quality
                  coverage: 11.81x in Q20 bases; agarose-fp
                  -----
                  * NOTE: This is a 'working draft' sequence. It currently
                  * consists of 3 contigs. The true order of the pieces
                  * is not known and their order in this sequence record is
                  * arbitrary. Gaps between the contigs are represented as
                  * runs of N, but the exact sizes of the gaps are unknown.
                  * This record will be updated with the finished sequence
                  * as soon as it is available and the accession number will
                  * be preserved.
                  *
                  * 1 5005: contig of 5005 bp in length
                  * 5006 5105: gap of 100 bp
                  * 5106 7812: contig of 2707 bp in length
                  * 7813 7912: gap of 100 bp
                  * 7913 335799: contig of 327887 bp in length.
FEATURES           Location/Qualifiers
source            1..335799
                  /organism="Mus musculus"
                  /mol_type="genomic DNA"
                  /db_xref="taxon:10090"
                  /chromosome="X"
                  /clone_lib="RPCI-23"
                  /clone="RP23-469F12"
                  /notes="assembly_fragment:01025
                  fragment_chain:1"
misc_feature      5106..7812
                  /note="assembly_fragment:00821
                  fragment_chain:1"
misc_feature      7913..335799
                  /note="assembly_fragment:03321"
ORIGIN
Alignment Scores:

```

```

Pred. No.:      8.77e-113      Length:      335799
Score:          130.00          Matches:      130
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:    9.10%           Indels:      0
DB:             2              Gaps:        0

US-09-787-097-12 (1-1429) x BX005139 (1-335799)

QY 1259 AsnPheThrTrpProLleLysLleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278
Db 63839 AATTTCATTGGCCCATCAAAATTCAGATTGCCTTCTCCAGCAGCAGCAACTTCATGGAC 63898
QY 1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAa 1298
Db 63899 CTGGTACAGTCTTCGTGACTTTCTTCAGTGTTCCTCTCGTCTTCGTGGTGGTGA 63958
QY 1299 ValValTrpLysLleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArg 1318
Db 63959 GTGGTCTGGAAGATCAAGCAGAGCTGTGGCATCCAGCGGAGAGAGCAACTTCTCGG 64018
QY 1319 GluMetGlnGlnMetAlaSerArgPropheAlaSerValAlaLeuGluThrAsp 1338
Db 64019 GAGATGCAACAGATGGCCGCCCTTCTGTTCTGTAAACGTTCCTTGGAAACAGAT 64078
QY 1339 GluGluProProAspLeuLleGlyGlySerIleLysThrValProLysProLleAlaLeu 1358
Db 64079 GAAGAACTCTCTGATCTCATTGGGGAGATATAAGACCGTCTTAAAGCCATTGCCCTG 64138
QY 1359 GluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGly 1378
Db 64139 GAGCCCTGCTTTGGTAAACAAAGCGCAGTCTCTCTGTATTCTGAGGCTCCCTCGAGGA 64198
QY 1379 LeuGlyGlyLeuProProGlyGlnSer 1388
Db 64199 CTGGAGGAATCCCTCTCTCTGTCAGTCA 64228

RESULT 21
AF301008          315 bp   mRNA      linear   MAM 26-SEP-2000
LOCUS             Bos taurus attractin mRNA, partial cds.
DEFINITION        AF301008
ACCESSION         AF301008
VERSION           AF301008.1 GI:10304438
KEYWORDS          Bos taurus (cow)
SOURCE            Bos taurus
ORGANISM          Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
                  Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
                  Bovidae; Bovinae; Bos.
REFERENCE         1 (bases 1 to 315).
                  Glad,T.W., Granholm,N.H., Westby,C.A., Marshall,D.M. and Kraft,E.
                  Bovine mahogany/attractin cDNA, including exons 26-29
                  Unpublished
JOURNAL           JOURNAL
AUTHORS           Glad,T.W., Granholm,N.H., Westby,C.A., Marshall,D.M. and Kraft,E.
TITLE             Direct Submission
JOURNAL           Submitted (29-AUG-2000) Bio/Micro, South Dakota State University,
                  2140D, 252 NPB, Brookings, SD 57007, USA
FEATURES           Location/Qualifiers
source            1..315
                  /organism="Bos taurus"
                  /mol_type="mRNA"
                  /db_xref="taxon:9913"
                  /tissue_type="cerebrum"
                  /notes="mahogany"
                  /codon_start=2
                  /product="attractin"
                  /protein_id="AAGI6254.1"
                  /db_xref="GI:10304439"
                  /translation="HSNFVDLVQFFVTFPFCFLSLLLVAVVKIKQSCWASRRRLQ
                  LRMQOMASRPFAVNVVALETDEPDLIGGSIKTVPKPIALEPCFGKNAVLVFR
                  LPR"
CDS

```

```
ORIGIN
Alignment Scores:
Pred. No.: 3-53e-90 Length: 315
Score: 104.00 Matches: 104
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 7.28% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AF301008 (1-315)
Qy 1273 HisSerAspPheMetAspLeuValGlnPheValThrPhePheSerCysPheLeuSer 1292
Db 2 CACAGCACTTCATGACCTGATACAGTTCCTCGGACCTTCCTCAGTTGTTCTCTCT 61

Qy 1293 LeuLeuValAlaAlaValValTrrPlySileLysGlnSerCysTrrPAlaSerArgArg 1312
Db 62 CTGCTCCTGGTGGCTGCTGCTGTTGGAGATCAACAAGTTGTTGGCTTCCAGAGCT 121

Qy 1313 ArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsn 1332
Db 122 AGAGAGCAACTTCTTCGAGAGATGCGAGATGGCCAGTCGTCCTTTGCTCTCTAAAT 181

Qy 1333 ValAlaLeuGluThrAspGluProProAspLeuIleGlySerIleLysThrVal 1352
Db 182 GTCGCTTGGAAACAGATGAAGACCTCTCTGATCTATTGGGGGAGATATAAGACTGTT 241

Qy 1353 ProLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPhe 1372
Db 242 CCCAGCCCATTTGCCCTGGAGCGGTGTTTGGCAACAAGCTGCTCTCTCTGTGTT 301

Qy 1373 ValArgLeuPro 1376
Db 302 GTGAGGCTCCCT 313

RESULT 22
LOCUS AR164813 2419 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 8 from patent US 6274339.
ACCESSION AR164813
VERSION AR164813.1 GI:16238022
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 2419)
AUTHORS Moore,K. and Magle,D.Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body weight disorders, including obesity
JOURNAL Patent: US 6274339-A 8 14-AUG-2001;
FEATURES
source
1..2419
Location/Qualifiers
/mol_type="unknown"
/organism="unknown"
ORIGIN
Alignment Scores:
Pred. No.: 7-67e-80 Length: 2419
Score: 94.00 Matches: 94
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.58% Indels: 0
DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x AR164813 (1-2419)
Qy 513 AlaLeuTyrValHisGlyTrrPlySAlaPheSerAlaAsnLysTrrArgLeuAlaAsp 532
Db 1803 GCTCTGTACGTTTCATGGTGGCTACAAAGGCTTTTCAGCGCAACAATACCGGCTTCAGAT 1862

Qy 533 AspLeuTyrArgTrrAspValAspThrGlnMetTrrThrIleLeuLysAspSerArgPhe 552
Db 533 AspLeuTyrArgTrrAspValAspThrGlnMetTrrThrIleLeuLysAspSerArgPhe 552

Db 1863 GACCTCTACAGATACAGATGGGATACATCAGATGTGGACCATCTTCTTAAGGACACCCGATT 1922
Qy 553 PheArgTrrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572
Db 1923 TTCGTTTACTTGCATACAGCTGTGATAGTGAAGAACCATGCTGTGTTTGGAGGGAAC 1982
Qy 573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAla 592
Db 1983 ACACAAATGACACTTCCATGAGCCACGGTGCACAAATGCTTCTCTGGGACTTCATGGCT 2042
Qy 593 TyrAspIleAlaCysAspArgTrrSerValLeuProArgPro 606
Db 2043 TATGACATTGCTTGTGACCGATGGTCAAGTCTTCCCAGACCT 2084

RESULT 23
LOCUS E26747 1012 bp DNA linear PAT 18-JUN-2001
DEFINITION Structure and function of novel biorhythm marker gene.
ACCESSION E26747
VERSION E26747.1 GI:13026325
KEYWORDS JP 1999169185-A/1.
SOURCE Rattus sp.
ORGANISM Rattus sp.
REFERENCE
1 (bases 1 to 1012)
AUTHORS Mario,i., Yutaka,S., Satoru,S. and Mikio,K.
TITLE Structure and function of novel biorhythm marker gene
JOURNAL Patent: JP 1999169185-A 1 29-JUN-1999;
AGENCY OF IND SCIENCE & TECHNOL,OTSUKA PHARMACEUT CO LTD
COMMENT
OS Rattus sp. (Wister rat)
PN JP 1999169185-A/1
PD 29-JUN-1999
PF 12-DEC-1997 JP 1997362890
PR
PI MARIO ISHIDA,YUTAKA SADAKANE,SATORU SUZUKI,MIKIO KIKUCHI PC
C12N15/09,C07K14/47,C12Q1/68//C12N5/10,C12P21/02,(C12N15/09, PC
C12R1.91),
PC (C12N5/10,C12R1.91),(C12P21/02,C12R1.91),(C12N5/00,C12N5/00,
PC (C12N15/00,C12R1.91),(C12N5/00,C12R1.91)
CC Strandedness: Single;
CC Topology: Linear;
FH Key Location/Qualifiers
FT CDS
1..1012
Location/Qualifiers
/mol_type="Rattus sp."
/organism="genomic DNA"
/db_xref="taxon:10118"
ORIGIN
Alignment Scores:
Pred. No.: 3-21e-78 Length: 1012
Score: 92.00 Matches: 92
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 6.44% Indels: 0
DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x E26747 (1-1012)
Qy 849 SerLysLeuThrLeuThrProTrrPValGlyLeuArgLysIleAsnValSerTrrPys 869
Db 313 TCCAAGCTCACTCTGACTCCATGGTGGTCTTCCGGAAGATCAATGATCTTACTGTGC 372
Qy 869 TrrGluAspMetSerProPheThrAsnSerLeuLeuGlnTrrMetProSerGluProSer 888
Db 373 TGGGAGATATGCTCTCCATTCACAAATAGTTTGTGTCAGTGGATGCCATCTGAGCCAGT 432
Qy 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
Db 433 GATCGCGCTTCTGTGGATCTTGTGAGAGCTAGTACTCGGGGACTGAAGGCTGCAACC 492
```

```

QY 909 CysIleAsnProLeuAnGlySerValCysGluArgProAlaAsnHisSerAlaLysGln 928
Db 493 TGCATCAACCACTCAATGAAGTGTGTGAAGGCCCTGCAAAACACACAGTCCCAAGCAG 552

QY 929 CysArgThrProCysAlaLeuArgThrAlaCysGly 940
Db 553 TGGCGGACACCCCTGTGCCCTGCGGACAGCATGTGGT 588

RESULT 24
AB038388
LOCUS AB038388 4500 bp mRNA linear ROD 18-JAN-2001
DEFINITION Rattus norvegicus Atrn mRNA for attractin, complete cds.
ACCESSION AB038388
VERSION AB038388.1 GI:12275311
KEYWORDS attractin.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (sites)
AUTHORS Kuranoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T.,
Kawaguchi, S., Ogawa, Y., Nakao, K., Baren, G.S., Nagao, M., Ushijima, T.
and Serikawa, T.
TITLE Attractin/mahogan/zitter plays a critical role in myelination of
the central nervous system
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
MEDLINE 21143347
PubMed 11209055
REFERENCE 2 (bases 1 to 4500)
AUTHORS Kuranoto, T., Serikawa, T. and Ushijima, T.
TITLE Direct Submission
JOURNAL Submitted (15-FEB-2000) Takashi Kuranoto, National Cancer Center
Research Institute, Carcinogenesis Division, Tsukiji 5-1-1,
Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuranoto@ncc.go.jp,
Tel:81-33542-2511(ex.4521), Fax:81-35565-1753)
FEATURES
source
1..4500
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="WTC"
/db_xref="taxon:10116"
1..4500
/gene="Atrn"
/gene="Atrn"
/codon_start=1
/product="attractin"
/protein_id="BAB21018.1"
/db_xref="GI:12275312"
/translation="MVAA...A...E...R...L...G...T...T...A...P...A...G...K...H...Q...R...P...C...A...T...G...A...N...R...P...
WRAGLCLRVSRALSPPLPLPLPLLSLLLEPREAAVA...A...V...S...G...A...A...B...A...K...
EDRPGVNGRCNPGTCGCTGWMVQCQCGFRITGSGFVTPDGNKYKTK
CTWLEIQPNKIMRLFNHFATFECSDWHLVYDGSIVAPLIAAFSGLIVPERDNET
APRVTVSGVAILHPDSDAVNLGTNITNTDMPNCSGRGCKSNSTVSEC
SENWGESCDIPEHCTDCCGPHRGICNADTRGCSGCFHWQPGGCIIPVANOSEFWR
EYFSDKLPRAHKAENVNIMVVGGMFNSHSDYSMLAYDLASREWLNLHNSVSV
VYRIGLSLALHKDKITWYGGKIDSTGNVTNEARVPHHNSWLLTPAKQOYAVVGH
SAHIVTSGRVWYVIFGCELYGISVYQVYDLEKNTWSILQTOGALVGGYHSS
VDYDRTKALYVHGXYKAFSANKYRLADDLVYHVDTPQWMTILKDSFRFRYLTAIVIS
GTMVFGVGNTHNDTSMHGAKCFSSDFMAYDIACDRWSVLPRLPHLDVNRFGHSAVL
HNSTWYFGLNSCPPLNSMCLBDSRVNWSFIFHPACQNGHSCINQICERKEDLTT
LATEOAEKLKSECFSKRTLDDHRCDDQHTDCYCTANTNDCHWCDHCVPVNHSCTEG
QTSIAXIDNCPDNPMYCNKITSCESSCALDONQWEPRNOECIALPENIGIGHLY
GNSCLKITAKENYDNKUSCRNHPASLASLSQKVEFVLKQLMSSQSTSKLTL
TPWVGLRKINYNWDMSPFTNSLLQWSPSDAGFGLLSBPSRGLKAACTCIN
PLNGSCERPANHSAQCTPCALRTACGECTSSSECMWCSNMKQCVDSNAYVASFP
FGOCMEWYTWSSCPPECSCGYCTCSHLEQPGCGWCTDPTSNITGKGCIEGYSYKPVQM
PASHAGNVYPOPPLNSMCLBDSRVNWSFIFHPACQNGHSCINQICERKEDLTT
GKHCFNCISGFDPTNGKCPCKCNKGNHSLCNTNKGCTCTTGVGEGECOLCEVE
NRYQGNPLKGTCTCYTLLIYQTFSLSQEDDRYTAIVNVPATPDQRNDLQMPFNASK
NFNLNITWATSFPAGTQGTGEVFSVSKINIKYKDSFNEKDFDRNHNFFVTVSN

```

```

ORIGIN
Alignment Scores: 1.06e-77 Length: 4500
Pred. No.: 92.00 Matches: 353
Score: 92.00 Conservative: 0
Percent Similarity: 95.15% Mismatches: 9
Best Local Similarity: 95.15% Indels: 18
Query Match: 6.44% Gaps: 0
DB: 10

US-09-787-097-12 (1-1429) x AB038388 (1-4500)

QY 849 SerLysLeuThrLeuThrProTrrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
Db 2614 TCCAGCTCACTCTGACTCCATGGTGGTCTTCGGAAGATCAATGTATCTTACTGTGTC 2673

QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuThrMetProSerGluProSer 888
Db 2674 TGGAGGATATGCTCCATTCACAAATAGTTTGTCTGAGTGGATGCCATCTGAGCCAGT 2733

QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
Db 2734 GATCCCGCTTCTGTGGATCTTGTACAGCCTAGTACTCGGGGACTGAAGGCTGCAACC 2793

QY 909 CysIleAsnProLeuAnGlySerValCysGluArgProAlaAsnHisSerAlaLysGln 928
Db 2794 TGCATCAACCACTCAATGAAGTGTGTGCAAGGCCCTGCAAAACACACAGTCCCAAGCAG 2853

QY 929 CysArgThrProCysAlaLeuArgThrAlaCysGlyAsp-CysThrSerGly-SerSerG 948
Db 2854 TGGCGGACACCCCTGTGCCCTGCGGACAGCATGTGTGA-GTGCACCCAGCAG-CAGTCCG 2911

QY 948 luCysMetTrpCysSerAsnMethLysGlnCysValAspSerAsnAlaTyrValAlaSerP 968
Db 2912 AATGCATGTGTGTGAGCAGCAACATGAAGCAGTGCCTAGTACTCAATGCCCTAGTACTCT 2971

QY 968 heProPheGlyGlnCysMetGluTrpTyrThrMetSerThr-CysProProGluAsnCys 987
Db 2972 TTCCTTTTCGCCCACTGTATGGAGTGGTATACAATGAGCAG-CTGCCCACTGAAAATGTC 3030

QY 988 SerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTyrCysThrAsp 1007
Db 3031 TCTGGCTACTGTACTCTGACGCCATTCCTTGGAGCAGCCAGGCTGTGTTGGTGTACTGAT 3090

QY 1008 ProSerAsnThrGlyLysGlyCysIleGluGlySerTyrLysGlyProValLysMet 1027
Db 3091 CCTAGCAATATCTGGGAAAGGCAATGTATTGAGGGAAGCTATAAGGACCCGCTGAAGATG 3150

QY 1028 ProSerGlnAlaPro--ThrGlyAsnPhe-TyrProGlnProLeuLeuAsnSerSerMet 1046
Db 3151 CCTTCACA-TG-CCTCTACAGGAAATGT-CTACCCACAGGCCCTTCTGAATCTCAGCATG 3207

QY 1047 CysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysAsn 1066
Db 3208 TGTCTAGAGACACAGACATCAACTGGTCTTTTATTCACTGTGCTGAGTTCCTCCAGTGCAT 3267

QY 1067 GlyHisSerLysCysIleAsnGlnSerLysCysGluLysCysGluAsn-LeuThrThrGly 1086
Db 3268 GGACACAGCAAGTGCATCAACAGAGCATCTGTGAGNAGTGTGAGGA-CCTGACCCACAGG 3326

QY 1086 YLYeHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCys 1106
Db 3327 CAAGCACTGGAGACCTGCATATCTGGCTCTATGTGTGACCCGCACTAATGGAGGCAAAATG 3386

QY 1106 sGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPh 1126
Db 3387 TCAGCCATGCAAGTGCACCGGACCGCATCACTGTGTGCAACCAACCACTGGCAAGTGCCT 3446

QY 1126 eCysThrThrLysGlyValLysGlyAsp-GluCysGlnLeuLeuCysGluValGluAsnArg 1146
Db 3447 CTGTACCAACCAAGGCGCTCAAGGAGGA-GGAGTGCCAGCTATGTGAGGTAGAAAAATCGAT 3505

```

```
QY 1146 yrglnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuLeuLeuAspTyrGlnPhe 1165
Db 3506 ACCAGGGAACCTCTCAA-AGGAACATGTTACTATACCTTCTCATTCAGTATCAGTTC 3564
QY 1166 ThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheValAlaThr 1185
Db 3565 ACTTTTGGCTGCTCCAGGAAGACACCGCTATTACAGCCATCAATTTTGGCTACT 3624
QY 1186 ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu 1205
Db 3625 CCTGATGAACAAACAGGAGATTGGACATGTTTCATCAATGCTCCAAACAACTTCAACCTC 3684
QY 1206 AsnIleThrTrpAla 1210
Db 3685 AACATCACCTGGGCC 3699

RESULT 25
AR273223/c
LOCUS AR273223 246 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 966 from patent US 6504010.
ACCESSION AR273223
VERSION AR273223.1 GI:29705108
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 4.38e-69 Length: 246
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.74% Indels: 0
DB: Gaps: 0
US-09-787-097-12 (1-1429) x AR273223 (1-246)
QY 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 187
QY 859 LeuArgLysIleAsnValSerTyrTyrCysTyrTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTTCGGAAGATCAATGTCTCTACTGTCTGCTGGAGAGATATGTCCTCCATTTACAAATAGT 127
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACAGTGGATGCGCTCTGAGCCAGTCAGTGTGGATTCTGTGGAATTTTATCAGAA 67
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCAGTACTCGGGAGTGAAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 919 GluArg 920
Db 6 GAAAGG 1

US-09-787-097-12 (1-1429) x AR273223 (1-246)
QY 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 187
QY 859 LeuArgLysIleAsnValSerTyrTyrCysTyrTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTTCGGAAGATCAATGTCTCTACTGTCTGCTGGAGAGATATGTCCTCCATTTACAAATAGT 127
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACAGTGGATGCGCTCTGAGCCAGTCAGTGTGGATTCTGTGGAATTTTATCAGAA 67
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCAGTACTCGGGAGTGAAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 919 GluArg 920
Db 6 GAAAGG 1

RESULT 26
AR276804/c
LOCUS AR276804 246 bp DNA linear PAT 10-APR-2003
DEFINITION Sequence 966 from patent US 6509448.
ACCESSION AR276804
```

```
VERSION AR276804.1 GI:29710451
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 4.38e-69 Length: 246
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.74% Indels: 0
DB: Gaps: 0
US-09-787-097-12 (1-1429) x AR276804 (1-246)
QY 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCCAAAGCTCACCTTAACCCCATGGTGGC 187
QY 859 LeuArgLysIleAsnValSerTyrTyrCysTyrTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTTCGGAAGATCAATGTCTCTACTGTCTGCTGGAGAGATATGTCCTCCATTTACAAATAGT 127
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACAGTGGATGCGCTCTGAGCCAGTCAGTGTGGATTCTGTGGAATTTTATCAGAA 67
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCAGTACTCGGGAGTGAAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 919 GluArg 920
Db 6 GAAAGG 1

RESULT 27
AR407079/c
LOCUS AR407079 246 bp DNA linear PAT 18-DEC-2003
DEFINITION Sequence 966 from patent US 6630574.
ACCESSION AR407079
VERSION AR407079.1 GI:40156890
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
ORIGIN
Alignment Scores:
Pred. No.: 4.38e-69 Length: 246
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservativeness: 0
US-09-787-097-12 (1-1429) x AR276804 (1-246)
```

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.74% Indels: 0  
DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x AR407079 (1-246)

QY 839 LeuArgileMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858  
DB 246 CTCGGATATGAGTCATCTCAGACATGTCACAGCTCACCTTAACCCATGGGTGGC 187  
QY 859 LeuArgileMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858  
DB 186 CTCGGATATGAGTCATCTCAGACATGTCACAGCTCACCTTAACCCATGGGTGGC 127  
QY 879 LeuLeuGlnThrMetProSerGluProSerAspAlaGlyPheCysGlyLeuLeuSerGlu 898  
DB 126 TTTACTACAGTGGATGCGGTCTGAGCCAGTGAATCTGGATTTCTGGAAATTTTATCAGAA 57  
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918  
DB 66 CCCAGTACTCGGGAGTGAAGGCTGCACTGCATCAACCCACTCAATGGTAGTCTGT 7  
QY 919 GluArg 920  
DB 6 GAAAGG 1

## RESULT 28

AX368256/c  
LOCUS AX368256 246 bp DNA linear PAT 16-FEB-2002  
DEFINITION Sequence 966 from Patent WO204514.  
ACCESSION AX368256  
VERSION AX368256.1 GI:18856329

KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Homo sapiens  
REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1

TITLE Wang, T., Watanabe, Y., Henderson, R.A., Johnson, J.C., Retter, M.W.,  
Marnett, K.S., Carter, D., Fanger, G.R., Vedvick, T.S., Bangur, C.S.,  
McNabb, A., Fanger, N., Switzer, A., McNeill, P.D. and Clapper, J.D.  
Compositions and methods for the therapy and diagnosis of lung  
cancer

JOURNAL Patent: WO 0204514-A 966 17-JAN-2002;  
CORIXA CORPORATION (US)

## FEATURES

source  
1..246  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Alignment Scores:  
Pred. No.: 4.38e-69 Length: 246  
Score: 82.00 Matches: 82  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.74% Indels: 0  
DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x AX368256 (1-246)

QY 839 LeuArgileMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858  
DB 246 CTCGGATATGAGTCATCTCAGACATGTCACAGCTCACCTTAACCCATGGGTGGC 187  
QY 859 LeuArgileMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858  
DB 186 CTCGGATATGAGTCATCTCAGACATGTCACAGCTCACCTTAACCCATGGGTGGC 127  
QY 879 LeuLeuGlnThrMetProSerGluProSerAspAlaGlyPheCysGlyLeuLeuSerGlu 898  
DB 126 TTTACTACAGTGGATGCGGTCTGAGCCAGTGAATCTGGATTTCTGGAAATTTTATCAGAA 57

QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918  
DB 66 CCCAGTACTCGGGAGTGAAGGCTGCACTGCATCAACCCACTCAATGGTAGTCTGT 7  
QY 919 GluArg 920  
DB 6 GAAAGG 1

## RESULT 29

AL954766\_2

## WPCOMMENT

Sequence split into 4 fragments LOCUS AL954766 Accession AL954766

Fragment Name	Begin	End
AL954766_1	1	110000
AL954766_2	100001	210000
AL954766_3	200001	310000
AL954766_4	300001	358753

Continuation (3 of 4) of AL954766 from base 200001 (AL954766 Mus musculus chromosome X c

Alignment Scores:  
Pred. No.: 4.51e-64 Length: 110000  
Score: 79.00 Matches: 79  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.53% Indels: 0  
DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x AL954766\_2 (1-110000)

QY 1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278  
DB 60740 AATTTCATCTGGCCCATCAAAATTCAGATTCCTTCCAGCAGCACTTCATGGAC 60799  
QY 1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAla 1298  
DB 60800 CTGGTACAGTCTTCGTGACTTTCTTCAGTTGTTCTCTCTGCTGCTTCGTGGCTGCA 60859  
QY 1299 ValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgArgGlnGlnLeuArg 1318  
DB 60860 GTGCTCTGGAGATCAAGCAGAGCTGTGGCATCAGCGGGAGAGCACTTCCTCGG 60919  
QY 1319 GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337  
DB 60920 GAGATGCAACAGATGGCCGCGCCCTTGTCTCTTAACCTTGCCTTGGAAACG 60976

## RESULT 30

BX004762\_3

## WPCOMMENT

Sequence split into 5 fragments LOCUS BX004762 Accession BX004762

Fragment Name	Begin	End
BX004762_1	1	110000
BX004762_2	100001	210000
BX004762_3	200001	310000
BX004762_4	300001	410000
BX004762_5	400001	442165

Continuation (4 of 5) of BX004762 from base 300001 (BX004762 Mus musculus chromosome X c

Alignment Scores:  
Pred. No.: 4.51e-64 Length: 110000  
Score: 79.00 Matches: 79  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.53% Indels: 0  
DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x BX004762\_3 (1-110000)

QY 1259 AsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAsp 1278  
DB 46152 AATTTCATCTGGCCCATCAAAATTCAGATTCCTTCCAGCAGCACTTCATGGAC 46211  
QY 1279 LeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAla 1298

```

Db      46212  CTGGTACAGTTCTTCTGACTTTCTTCTAGTTCTTCTCTGCTCTCTCTGCTGCA 46271
QY      1299  ValValTrpLysLysGlnSerCysTrpAlaSerArgArgArgGlnGlnLeuLeuArg 1318
Db      46272  GTGGTCTGAAGATCAAGCAGAGCTGTGTGGCATCCAGCGGAGAGCACTTCTCGG 46331
QY      1319  GluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337
Db      46332  GAGATGCAACAGATGCCAGCCGCCCTTGTCTCTGTAACCTTGCCCTTGGAAACG 46388

RESULT 31
LOCUS   H009384S15
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exons 17 and 18.
ACCESSION AF218903
VERSION   AF218903.1 GI:8118069
KEYWORDS
SEGMENT
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

REFERENCE
AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.
and Duke-Cohan,J.S.
TITLE     Secreted and membrane attractin result from alternative splicing of
the human ATRN gene
JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE   20283944
PUBMED    10811918
REFERENCE
AUTHORS   Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,
Schlossman,S.F. and Duke-Cohan,J.S.
TITLE     Direct Submission
JOURNAL   Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,
Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,
USA

FEATURES
Source    Location/Qualifiers
1..3257
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
1441..1626
/gene="ATRN"
/number=17
2189..2422
/gene="ATRN"
/number=18

exon
exon

ORIGIN
Alignment Scores:
Pred. No.: 2,26e-63 Length: 3257
Score: 77.00 Matches: 77
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.39% Indels: 0
DB: Gaps: 0

US-09-787-097-12 (1-1429) x H009384S15 (1-3257)

QY      985  GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyTrp 1004
Db      2191  GAAATTTGTCAGGCTACTGTACTCTGTACTGTCTGTCGACCAACAGGCTGTGGCTGG 2250
QY      1005  CysThrAspProSerAsnThrGlyLysGlyLysCysValleGlySerTyrLysGlyPro 1024
Db      2251  TGTACTGATCCGACGATACTGCGCAAGGGAATGCATAGAGGGTTCCTATAAGGACCA 2310
QY      1025  ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044

```

```

Db      2311  GTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAATCC 2370
QY      1045  SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061
Db      2371  AGCATGTCTTAGAGGACAGCAGATACAACTGGTCTTTTCATTCACTGCCA 2421

RESULT 32
LOCUS   HSDJ741H3
DEFINITION Human DNA sequence from clone RP4-741H3 on chromosome 20. Contains
the 3' end of the ATRN gene for attractin (with dipeptidylpeptidase
IV activity) (KIAA0548), ESTs, STSS and GSSs, complete sequence.
AL132773
ACCESSION AL132773.14 GI:7159777
VERSION   AL132773.14
KEYWORDS  HTG: ATRN; attractin; KIAA0548.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens

REFERENCE
1 (bases 1 to 104907)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Sehra,H.
Direct Submission
Submitted (19-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Mar 6, 2000 this sequence version replaced gi:7007301.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 20, constructed by the Sanger Centre Chromosome 20
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr20
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. RP4-741H3 is from the
library RPCI-4 constructed by the group of Pieter de Jong. For
further details see
http://www.choiri.org/bacpac/home.htm
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP4-741H3 The true
right end of clone RP4-581P3 is at 79521 in this sequence.

FEATURES
Source    Location/Qualifiers
1..104907
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clones="RP4-741H3"
/clone_lib="RPCI-4"
2..103785
/gene="ATRN"
join(<2..161,1842..2010,12065..12155,13334..13577,
14948..15131,15681..16035,18020..18104,25403..25623,
28499..28620,29531..29669,31254..31444,34807..35026,
36570..36755,37319..37552,43841..43978,47151..47244,
49049..49101,50578..50672,53618..53696,56777..56934,
77183..77244,86583..87061,91500..91607,96826..96940,
99395..103785)
/gene="ATRN"

```



IV activity) membrane bound isoform (X1AA0548)"))  
 /notes="match: cDNA: Em:AF116897 Em:AF119821 Em:AF106961  
 Em:AF034957 Em:AB011106 Em:AK000356 Em:AB011120  
 match: ESTs: Em:R87660 Em:H11351 Em:AA350292 Em:T69367  
 Em:AA368140 Em:A1372811 Em:AW141339 Em:AL048842"  
 /evidence=not\_experimental  
 join(c2..161,1842..2010,12065..12155,13334..13577,  
 14948..15131,15881..16035,18020..18104,25403..25623,  
 28499..28620,29531..29669,31254..31444,34807..35026,  
 36570..36755,37319..37552,43841..43978,47151..47244,  
 49049..49101,50578..50672,53618..53696,56777..56934,  
 7183..77244,86983..87061,91500..91607,96826..96940,  
 9395..93519)  
 /genes="ATRN"  
 /notes="match: proteins: Tr:O60283 Tr:Q9WU60 Tr:Q9WU77  
 Tr:O60295 Tr:O95414 Tr:O75882 Sw:Q19981 Sw:P19137  
 Sw:P25391"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /product="dJ74IH3.1.2 (attractin (with dipeptidylpeptidase  
 IV activity) membrane bound isoform (X1AA0548))"  
 /protein\_id="CAC32457.1"  
 /db\_xref="GI:13160052"  
 /db\_xref="GOA:O75882"  
 /db\_xref="SWISS-PROT:O75882"  
 /translation="ISNSDVTVECCSENWKGACDIPHTDNCQFPHRGICNSSDVR  
 GSCFSDWQPGCSVPVPAQSFWESEYSLKLPASHKAVNGINMVGGMFNH  
 SDYMWLADLAREMLPNRSVNNVVRIGHSLALYKDKIYMGCKIDSGNVTNEL  
 RVFHINESWLLTPKAGQYAVGSHAVITLKNRGVVMVIFGHCPLYGISNVQE  
 YLDKNTWILHTQALQVGGYSHVDHRTALYVGGYKAFSANKYRADDLYRY  
 DVTQMTLTKDSRFFRYLHTAVISGTMVFGNTHDTSNMGAKCFSDFWAYDI  
 ACDRWSLRPDLHDVNEFGHSAVLHNSMTVMVFGFNSLLSDILFTSECDAAHS  
 EAACLAAGGIRCVWNTGSSQISWALATDEQEEKLKSCEFSKRLDHRCDQHTDCY  
 SCTANTNDCHWNCDCVPRNHSCEQISIFRYENCNDKNPMYCNKTKSCRSALQD  
 NCQWEPNCEIAPLNICIGWHLVGNSECLIKITTAKENYDNAKLPFCRNHALLASLT  
 TQKKVEFLVQLRIMGSSQSMKLTLPWVGLRKINVSVCWEDMSPTNLLQWMP  
 EPSDAFCGILSEPTSTGLKAATCINPLNGSCVCRPANHSKQCTPCALRTACGDT  
 SGSECCWNNKQCVDSNAYVAFPGQCEWYMTSTCPENCSCYCTGCHCLQPG  
 CGMTDPSNTGKGCIGSGYKGVKMPQAPTFNFPQPLNNSMCLSDRYNWSFIH  
 CPACQCNHSGKICNOSICEKCNLTGKHCECTISGFYDPTNGCKQPCKNCHASL  
 CYNNTGKCTCTGKVGDCOLCEVENRYOQNPLSGTCYTYLLIDYQFTFSLQEDDR  
 YTAINFVATPEQNDLDMFINASKNFNLNITWAAFSAGTQAGEEMPVSVKTIKE  
 YKDSFNEKFDPRHNITFFVVSNTFWPKIQAFOHSNFMOLVQFFVTFSCFL  
 SILLVAVVKIKQSWASRRERQLREMQQASFPASVVALETDEEPLDGGSI  
 KTVPKPIALEPCFGKAAVLVSFVRLPRLGLGIPPPGOSGLAVASALVDISQMPDIV  
 KEKSGAVNRKQPPAQPCTCI"  
 join(c2..161,1842..2010,12065..12155,13334..13577,  
 14948..15131,15881..16035,18020..18104,25403..25623,  
 28499..28620,29531..29669,31254..31444,34807..35026,  
 36570..36755,37319..37552,43841..43978,47151..47244,  
 49049..49101,50578..50672,53618..53696,56777..56934,  
 68500..68517)  
 /genes="ATRN"  
 /notes="match: proteins: Tr:O75882 Tr:Q9NZ58 Tr:O95414"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /product="dJ74IH3.1.1 (attractin (with dipeptidylpeptidase  
 IV activity) secreted isoform)"  
 /protein\_id="CAC32456.1"  
 /db\_xref="GI:13160051"  
 /db\_xref="GOA:O75882"  
 /db\_xref="SWISS-PROT:O75882"  
 /translation="ISNSDVTVECCSENWKGACDIPHTDNCQFPHRGICNSSDVR  
 GSCFSDWQPGCSVPVPAQSFWESEYSLKLPASHKAVNGINMVGGMFNH  
 SDYMWLADLAREMLPNRSVNNVVRIGHSLALYKDKIYMGCKIDSGNVTNEL  
 RVFHINESWLLTPKAGQYAVGSHAVITLKNRGVVMVIFGHCPLYGISNVQE  
 YLDKNTWILHTQALQVGGYSHVDHRTALYVGGYKAFSANKYRADDLYRY  
 DVTQMTLTKDSRFFRYLHTAVISGTMVFGNTHDTSNMGAKCFSDFWAYDI  
 ACDRWSLRPDLHDVNEFGHSAVLHNSMTVMVFGFNSLLSDILFTSQCDAHS  
 EAACLAAGGIRCVWNTGSSQISWALATDEQEEKLKSCEFSKRLDHRCDQHTDCY  
 SCTANTNDCHWNCDCVPRNHSCEQISIFRYENCNDKNPMYCNKTKSCRSALQD  
 NCQWEPNCEIAPLNICIGWHLVGNSECLIKITTAKENYDNAKLPFCRNHALLASLT  
 TQKKVEFLVQLRIMGSSQSMKLTLPWVGLRKINVSVCWEDMSPTNLLQWMP

EPDAGRCGILSEPTSTGLKAATCINPLNGSCVCRPANHSKQCTPCALRTACGDT  
 SGSECCWNNKQCVDSNAYVAFPGQCEWYMTSTCPENCSCYCTGCHCLQPG  
 CGMTDPSNTGKGCIGSGYKGVKMPQAPTFNFPQPLNNSMCLSDRYNWSFIH  
 CPACQCNHSGKICNOSICEKCNLTGKHCECTISGFYDPTNGCKQPCKNCHASL  
 CYNNTGKCTCTGKVGDCOLCEVENRYOQNPLSGTCYTYLLIDYQFTFSLQEDDR  
 YTAINFVATPEQNDLDMFINASKNFNLNITWAAFSAGTQAGEEMPVSVKTIKE  
 YKDSFNEKFDPRHNITFFVVSNTFWPKIQAFOHSNFMOLVQFFVTFSCFL  
 SILLVAVVKIKQSWASRRERQLREMQQASFPASVVALETDEEPLDGGSI  
 KTVPKPIALEPCFGKAAVLVSFVRLPRLGLGIPPPGOSGLAVASALVDISQMPDIV  
 KEKSGAVNRKQPPAQPCTCI"  
 join(c2..161,1842..2010,12065..12155,13334..13577,  
 14948..15131,15881..16035,18020..18104,25403..25623,  
 28499..28620,29531..29669,31254..31444,34807..35026,  
 36570..36755,37319..37552,43841..43978,47151..47244,  
 49049..49101,50578..50672,53618..53696,56777..56934,  
 68500..68517)  
 /genes="ATRN"  
 /notes="match: STS: Em:AL110053 Em:HS74IH3S"  
 903..990  
 /note="MER2 repeat: matches 257..337 of consensus"  
 991..1297  
 /note="AluX repeat: matches 1..307 of consensus"  
 1299..1520  
 /note="MER2 repeat: matches 1..257 of consensus"  
 2718..2942  
 /note="AluS repeat: matches 1..225 of consensus"  
 2943..3127  
 /note="L1P repeat: matches 4678..4858 of consensus"  
 complement(3203..3804)  
 /note="match: GSS: Em:AQ381369"  
 3560..3672  
 /note="I2 repeat: matches 216..330 of consensus"  
 3881..3977  
 /note="MER81 repeat: matches 2..110 of consensus"  
 4326..4513  
 /note="I2 repeat: matches 738..924 of consensus"  
 4539..4682  
 /note="MER3 repeat: matches 1..149 of consensus"  
 4683..5195  
 /note="I2 repeat: matches 1383..1936 of consensus"  
 5209..5762  
 /note="MLR1E repeat: matches 2..562 of consensus"  
 5770..5835  
 /note="I2 repeat: matches 1949..2024 of consensus"  
 6074..6708  
 /note="I2 repeat: matches 2025..2732 of consensus"  
 6998..7305  
 /note="AluY repeat: matches 1..309 of consensus"  
 7663..7974  
 /note="AluY repeat: matches 1..311 of consensus"  
 complement(9805..10470)  
 /note="match: GSS: Em:AQ484998"  
 9825..10261  
 /gene="ATRN"  
 /note="match: GSS: Em:AQ338818"  
 9963..10167  
 /note="HALL repeat: matches 618..819 of consensus"  
 10173..10450  
 /gene="ATRN"  
 /note="match: GSS: Em:AQ413104"  
 10902..11155  
 /note="L1ME3 repeat: matches 5637..5902 of consensus"  
 12706..12802  
 /note="L1ME2 repeat: matches 6075..6168 of consensus"  
 12803..13235  
 /note="L1MA4 repeat: matches 5584..6168 of consensus"  
 13713..14023  
 /note="AluX repeat: matches 3..311 of consensus"  
 14407..14697  
 /note="AluJo repeat: matches 1..293 of consensus"  
 104907

Alignment Scores:  
 Pred. No.: 3,65e-62 Length: 104907  
 Score: 77.00 Matches: 77  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.39% Indels: 0  
 DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x HSDJ74IH3 (1-104907)





Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lenoczky, J., Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, J., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J., Meneses, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Oliver, J., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission

Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On May 12, 2000 this sequence version replaced gi:7582593.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## COMMENT

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WBIR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)

----- Project Information

Center project name: L5007

Center clone name: 561\_P\_16

----- Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 156504 bases at least Q40

Consensus quality: 161679 bases at least Q30

Consensus quality: 163964 bases at least Q20

Insert size: 170000; agarose-1p

Insert size: 165357; sum-of-contigs

Quality coverage: 4.5 in Q20 bases; agarose-1p

Quality coverage: 4.6 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 21 contigs. The true order of the pieces

\* is not known and their order in this sequence record is

\* arbitrary. Gaps between the contigs are represented as

\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence

\* as soon as it is available and the accession number will

\* be preserved.

1 1216: contig of 1216 bp in length

\* 1217 1316: gap of 100 bp

\* 1317 2985: contig of 1669 bp in length

\* 2986 3085: gap of 100 bp

\* 3086 4628: contig of 1543 bp in length

\* 4629 4728: gap of 100 bp

\* 4729 7507: contig of 2779 bp in length

\* 7508 7607: gap of 100 bp

\* 7608 11292: contig of 3685 bp in length

\* 11293 11392: gap of 100 bp

\* 11393 14213: contig of 2821 bp in length

\* 14214 14313: gap of 100 bp

\* 14314 17828: contig of 3515 bp in length

\* 17829 17928: gap of 100 bp

\* 17929 21068: contig of 3140 bp in length

\* 21069 21168: gap of 100 bp

\* 21169 25035: contig of 3867 bp in length

\* 25036 25135: gap of 100 bp

\* 25136 28331: contig of 3196 bp in length

\* 28332 28431: gap of 100 bp

\* 28432 33595: contig of 5164 bp in length

\* 33596 33695: gap of 100 bp

\* 33696 40346: contig of 6651 bp in length

\* 40347 40446: gap of 100 bp

\* 40447 43807: contig of 3361 bp in length

\* 43808 43907: gap of 100 bp

\* 43908 48827: contig of 4820 bp in length

\* 48828 56083: contig of 7256 bp in length

\* 56084 56183: gap of 100 bp

\* 56184 67065: contig of 10882 bp in length

\* 67066 67165: gap of 100 bp

\* 67166 82881: contig of 15716 bp in length

\* 82882 82981: gap of 100 bp

\* 82982 101845: contig of 18864 bp in length

\* 101846 101945: gap of 100 bp

\* 101946 122755: contig of 20810 bp in length

\* 122756 122855: gap of 100 bp

\* 122856 142457: contig of 19602 bp in length

\* 142458 142557: gap of 100 bp

\* 142558 167357: contig of 24800 bp in length.

FEATURES

source

1. .167357

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="20"

/map="20"

/clone="RP11-561P16"

/clone\_lib="RPC1-11 Human Male BAC"

1. .1216

/note="assembly\_fragment"

1317. 2985

/note="assembly\_fragment"

3086. 4628

/note="assembly\_fragment"

4729. 7507

/note="assembly\_fragment"

7608. 11292

/note="assembly\_fragment"

11393. 14213

/note="assembly\_fragment"

14314. 17828

/note="assembly\_fragment"

17929. 21068

/note="assembly\_fragment"

21169. 25035

/note="assembly\_fragment"

25136. 28331

/note="assembly\_fragment"

28432. 33595

/note="assembly\_fragment"

33696. 40346

/note="assembly\_fragment"

40447. 43807

/note="assembly\_fragment"

clone end:SP6

vector side:right

43908 -48727

/note="assembly\_fragment"

48828. 56083

/note="assembly\_fragment"

56184. 67065

/note="assembly\_fragment"

67166. 82881

/note="assembly\_fragment"

82982. 101845

/note="assembly\_fragment"

clone end:T7

vector side:right

101946. 122755

/note="assembly\_fragment"

122856. 142457

/note="assembly\_fragment"

142558. 167357

/note="assembly\_fragment"

## ORIGIN

Alignment Scores:  
 Pred. No.: 5,31e-62 Length: 167357  
 Score: 77.00 Matches: 77  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.39% Indels: 0  
 DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x AC055771 (1-167357)

QY 985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluInProGlyCysGlyTyr 1004  
 Db 113358 GAAATTTGTTTCAGGCTACTGCTAGCTAGTTCGTTGAGCAACAGGCTGTGGCTGG 113299  
 QY 1005 CysThrAspProSerAsnThrGlyLysGlyLysCysLeuGluGlySerTyrLysGlyPro 1024  
 Db 113298 TGACTGATCCCGAGCAATACTGGCAAGGGAATGCTAGAGGTTCTCTATAAGGACCA 113239  
 QY 1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044  
 Db 113238 GTGAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGGCCCTGCTCAATCC 113179  
 QY 1045 SerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCysPro 1061  
 Db 113178 AGCATGTGTAGGACAGCAGATACACTGGTCTTTTCATTCACTGTCCA 113128

## RESULT 35

AP002898 206647 bp DNA linear PRI 24-OCT-2000  
 LOCUS Homo sapiens genomic DNA, chromosome 20p, clone:13N6, complete  
 DEFINITION sequence.

ACCESSION AP002898

VERSION AP002898.1

GI:10998163

HTG.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 206647)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Homo sapiens 206,647 genomic DNA of 20p

Published Only in DataBase (2000)

2 (bases 1 to 206647)

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,

Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.

Direct Submission

Submitted (22-OCT-2000) Masahira Hattori, The Institute of Physical

and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

[E-mail:hattori@gscc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/]

Tel:81-45-503-9111, Fax:81-45-503-9170

Overlapping Clones: AL132773, AL109804

Cytogenetic Position: 20p

STS Markers: D20S828, D20S813, stSG408, A005005, sts-H22126,

D20S867, D20S752

Additional author information

Hayflick, S.J.

Oregon Health Sciences University, Molecular &amp; Medical Genetics

Dept. 3181 SW Sam Jackson Park Road, MC-L103, BSA-4596, Portland OR

97201-3098, USA

phone:1-503-494-6866, fax:1-503-494-6886

e-mail: hayflick@ohsu.edu

Location/Qualifiers

1..206647

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="20"

/map="20p"

/clone="13N6"

## ORIGIN

Alignment Scores:  
 Pred. No.: 6,29e-62 Length: 206647  
 Score: 77.00 Matches: 77  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.39% Indels: 0  
 DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x AP002898 (1-206647)

QY 985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluInProGlyCysGlyTyr 1004  
 Db 61935 GAAATTTGTTTCAGGCTACTGCTAGCTAGTTCGTTGAGCAACAGGCTGTGGCTGG 61994  
 QY 1005 CysThrAspProSerAsnThrGlyLysGlyLysCysLeuGlySerTyrLysGlyPro 1024  
 Db 61995 TGACTGATCCCGAGCAATACTGGCAAGGGAATGCTAGAGGTTCTCTATAAGGACCA 62054  
 QY 1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044  
 Db 62055 GTGAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCAGGCCCTGCTCAATCC 62114  
 QY 1045 SerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCysPro 1061  
 Db 62115 AGCATGTGTAGGACAGCAGATACACTGGTCTTTTCATTCACTGTCCA 62165

## RESULT 36

AX326809 207433 bp DNA linear PAT 07-JAN-2002  
 LOCUS Sequence 5 from Patent WO0178894.

ACCESSION AX326809

VERSION AX326809.1

GI:18097525

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 207433)

Keith,T.

Novel human gene relating to respiratory diseases, obesity, and

inflammatory bowel disease

Patent: WO 0178894-A 5 25-OCT-2001;

Genome Therapeutics Corp. (US)

Location/Qualifiers

1..207433

/organism="Homo sapiens"

/mol\_type="unassigned DNA"

/db\_xref="taxon:9606"

Alignment Scores:

Pred. No.: 6,31e-62 Length: 207433

Score: 77.00 Matches: 77

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 5.39% Indels: 0

DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x AX326809 (1-207433)

QY 985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluInProGlyCysGlyTyr 1004  
 Db 9815 GAAATTTGTTTCAGGCTACTGCTAGCTAGTTCGTTGAGCAACAGGCTGTGGCTGG 9874  
 QY 1005 CysThrAspProSerAsnThrGlyLysGlyLysCysLeuGlySerTyrLysGlyPro 1024  
 Db 9875 TGACTGATCCCGAGCAATACTGGCAAGGGAATGCTAGAGGTTCTCTATAAGGACCA 9934  
 QY 1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044

Db 9935 GTGAGATGCTTGGCAAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTCC 9994

QY 1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061  
 |||||  
 Db 9995 AGCATGTGTCTAGAGACAGCAGATACAACTGGCTTTTCATTCACCTGTCCA 10045  
 |||||

RESULT 37

AR164814  
 LOCUS AR164814 1051 bp DNA linear PAT 17-OCT-2001  
 DEFINITION Sequence 10 from patent US 6274339.  
 ACCESSION AR164814  
 VERSION AR164814.1 GI:16238023  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.

REFERENCE  
 1 (bases 1 to 1051)  
 Moore, K. and Nagle, D. Lynn.  
 Methods and compositions for the diagnosis and treatment of body  
 weight disorders, including obesity  
 TITLE Patent: US 6274339-A 10 14-AUG-2001;  
 JOURNAL Location/Qualifiers  
 FEATURES  
 source 1..1051  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Alignment Scores: Pred. No.: 8 39e-63 Length: 1051  
 Score: 76.00 Matches: 111  
 Percent Similarity: 98.23% Conservative: 0  
 Best Local Similarity: 98.23% Mismatches: 1  
 Query Match: 5.32% Indels: 2  
 DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x AR164814 (1-1051)

QY 87 AlaAlaValSerGlySerAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
 |||||

Db 463 GCGGCGCGGTGCTCCGGCTCGCGCGCAGCGGCGCAAGGAATGTACCGCGCGTGTCTC 522  
 |||||

QY 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys-ProAlaGlyTrpValG1 126  
 |||||

Db 523 AACGGCGCGCGTGCACCCCTGCGCGCGCGCGCTGCGCGCGCTGCGCGCGCTGCGCG 581  
 |||||

QY 126 yGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValTh 146  
 |||||

Db 582 CGAGCAATCCAGCACTGCGGGGCGGCTTCAGACTAACTGCTCTTCGGATTGTAAAC 641  
 |||||

QY 146 AspGlyProGlyAsnTyrGlyThrGlyThrGlyCysThrTrpLeuIleGluGlyGlnPr 166  
 |||||

Db 642 AGATGGACCTGGGAATTATAAATATAGACGAAGTGCACATGGCTCAATTGAAGGACAGCC 701  
 |||||

QY 166 oAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLe 186  
 |||||

Db 702 AATAGAAATAGAGACTTCGCTTCAACCATTTTGCTACAGATGTAGCTGGGACCAATT 761  
 |||||

QY 186 uTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198  
 |||||

Db 762 ATATGTTTATGATGGGAGCTCAATCTACGACCTCTG 798  
 |||||

RESULT 38

H009384811  
 LOCUS H009384811 3589 bp DNA linear PRI 14-DEC-2000  
 DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 12.  
 ACCESSION AF218899  
 VERSION AF218899.1 GI:8118065  
 KEYWORDS  
 SEGMENT 11 of 27  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 3589)  
 Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.  
 Secreted and membrane attractin result from alternative splicing of  
 the human ATRN gene  
 TITLE Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)  
 JOURNAL 20283944  
 MEDLINE 10811918  
 PUBMED 10811918

2 (bases 1 to 3589)  
 Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,  
 Schlossman, S.F. and Duke-Cohan, J.S.  
 Direct Submission  
 TITLE Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,  
 Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,  
 USA  
 FEATURES  
 Location/Qualifiers  
 1..3589  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20p13"  
 1585..1805  
 /gene="ATRN"  
 /number=12

Alignment Scores: Pred. No.: 1 89e-60 Length: 3589  
 Score: 74.00 Matches: 74  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.18% Indels: 0  
 DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x H009384811 (1-3589)

QY 624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPhe 643  
 |||||

Db 1583 AGCAACCATGTATGTTCGGTGGTTTCATAGTCTCTCTCAGCAGACATCTGTATTC 1642  
 |||||

QY 644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663  
 |||||

Db 1643 ACCTCGGACAGTGTGATGGCATCGAGTGAAGCCCTGTTTGTAGGAGGACCTGGT 1702  
 |||||

QY 664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683  
 |||||

Db 1703 ATTCTGGTGTGTGGACACAGAGGTCGTCTCAGTGTATCTCTGGGCGCTGGCAACTGAT 1762  
 |||||

QY 684 GluGlnGluGlyLysLeuLysSerGluCysPheSerLysArg 697  
 |||||

Db 1763 GAACAAGAGAAAAGTTAAATCAGAATGTTTTTCCAAAGA 1804  
 |||||

RESULT 39

AC023420/c  
 LOCUS AC023420 155869 bp DNA linear HTG 23-SEP-2000  
 DEFINITION Homo sapiens chromosome 2 clone RP11-737G10 map 2, WORKING DRAFT  
 SEQUENCE, 15 unordered pieces.  
 ACCESSION AC023420  
 VERSION AC023420.4 GI:10280876  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 155869)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 AUTHORS  
 TITLE Homo sapiens chromosome 2, clone RP11-737G10  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 155869)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Beda, F., Boguski, L.,  
 Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,

REFERENCE  
 1 (bases 1 to 3589)  
 Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.  
 Secreted and membrane attractin result from alternative splicing of  
 the human ATRN gene  
 TITLE Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)  
 JOURNAL 20283944  
 MEDLINE 10811918  
 PUBMED 10811918

2 (bases 1 to 3589)  
 Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,  
 Schlossman, S.F. and Duke-Cohan, J.S.  
 Direct Submission  
 TITLE Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,  
 Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,  
 USA  
 FEATURES  
 Location/Qualifiers  
 1..3589  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20p13"  
 1585..1805  
 /gene="ATRN"  
 /number=12

Alignment Scores: Pred. No.: 1 89e-60 Length: 3589  
 Score: 74.00 Matches: 74  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.18% Indels: 0  
 DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x H009384811 (1-3589)

QY 624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPhe 643  
 |||||

Db 1583 AGCAACCATGTATGTTCGGTGGTTTCATAGTCTCTCTCAGCAGACATCTGTATTC 1642  
 |||||

QY 644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663  
 |||||

Db 1643 ACCTCGGACAGTGTGATGGCATCGAGTGAAGCCCTGTTTGTAGGAGGACCTGGT 1702  
 |||||

QY 664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683  
 |||||

Db 1703 ATTCTGGTGTGTGGACACAGAGGTCGTCTCAGTGTATCTCTGGGCGCTGGCAACTGAT 1762  
 |||||

QY 684 GluGlnGluGlyLysLeuLysSerGluCysPheSerLysArg 697  
 |||||

Db 1763 GAACAAGAGAAAAGTTAAATCAGAATGTTTTTCCAAAGA 1804  
 |||||

RESULT 39

AC023420/c  
 LOCUS AC023420 155869 bp DNA linear HTG 23-SEP-2000  
 DEFINITION Homo sapiens chromosome 2 clone RP11-737G10 map 2, WORKING DRAFT  
 SEQUENCE, 15 unordered pieces.  
 ACCESSION AC023420  
 VERSION AC023420.4 GI:10280876  
 KEYWORDS HTG; HTGS PHASE1; HTGS DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 155869)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 AUTHORS  
 TITLE Homo sapiens chromosome 2, clone RP11-737G10  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 155869)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Beda, F., Boguski, L.,  
 Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,

Choepe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
DeArlano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,  
Fenster, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,  
Galan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., Landers, T., Lagouque, K., Lehoczy, J., Levine, R.,  
Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,  
McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J.,  
Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Naylor, J.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Olivat, T.M.,  
Peterson, K., Piers, N., Pisan, C., Pollara, V., Raymond, C.,  
Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,  
Severy, P., Spencer, B., Strange-Thomann, N., Stojanovic, N.,  
Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A.,  
Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,  
Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and  
Zody, M.

# TITLE JOURNAL

## COMMENT

Direct Submission  
Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Sep 23, 2000 this sequence version replaced gi:7767820.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RN/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence.submissions@genome.wi.mit.edu](mailto:sequence.submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: L6805  
Center clone name: 737 G 10  
----- Summary Statistics  
Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 145736 bases at least Q40  
Consensus quality: 151534 bases at least Q30  
Consensus quality: 153551 bases at least Q20  
Insert size: 157000; agarose-fp  
Quality coverage: 154469; sum-of-contigs  
Quality coverage: 4.2 in Q20 bases; agarose-fp  
Quality coverage: 4.3 in Q20 bases; sum-of-contigs  
-----

NOTE: This is a 'working draft' sequence. It currently  
consists of 15 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence,  
as soon as it is available and the accession number will  
be preserved.

1 12284: contig of 12284 bp in length  
\* 12285 12384: gap of 100 bp  
\* 12385 13704: contig of 1320 bp in length  
\* 13705 13805: gap of 100 bp  
\* 13805 15657: contig of 1853 bp in length  
\* 15658 15757: gap of 100 bp  
\* 15758 17755: contig of 1998 bp in length  
\* 17756 17855: gap of 100 bp  
\* 17856 19875: contig of 2020 bp in length  
\* 19876 19975: gap of 100 bp  
\* 19976 22894: contig of 2919 bp in length  
\* 22895 22994: gap of 100 bp  
\* 22995 26993: contig of 3998 bp in length  
\* 26993 27093: gap of 100 bp  
\* 27093 31423: contig of 4330 bp in length  
\* 31423 31523: gap of 100 bp  
\* 31523 37920: contig of 6398 bp in length  
\* 37921 38021: gap of 100 bp  
\* 38021 48644: contig of 10624 bp in length  
\* 48645 48744: gap of 100 bp

\* 48745 69112: contig of 20368 bp in length  
\* 69113 86214: gap of 100 bp  
\* 86214 86049: contig of 16837 bp in length  
\* 86050 86150: gap of 100 bp  
\* 86150 107477: contig of 21328 bp in length  
\* 107478 107577: gap of 100 bp  
\* 107578 131094: contig of 23517 bp in length  
\* 131095 131194: gap of 100 bp  
\* 131195 155869: contig of 24675 bp in length.  
FEATURES  
Location/Qualifiers  
source  
1. 155869  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="2"  
/map="2"  
/clone="RP11-737G10"  
/clone\_lib="RPC1-11 Human Male BAC"  
1. 12284  
/note="assembly\_fragment"  
clone end: SP6  
vector\_side:left  
12385..13704  
/note="assembly\_fragment"  
13805..15657  
/note="assembly\_fragment"  
15758..17755  
/note="assembly\_fragment"  
17856..19875  
/note="assembly\_fragment"  
19976..22894  
/note="assembly\_fragment"  
22995..26992  
/note="assembly\_fragment"  
27093..31422  
/note="assembly\_fragment"  
31523..37920  
/note="assembly\_fragment"  
38021..48644  
/note="assembly\_fragment"  
48745..69112  
/note="assembly\_fragment"  
69213..86049  
/note="assembly\_fragment"  
86150..107477  
/note="assembly\_fragment"  
107578..131094  
/note="assembly\_fragment"  
131195..155869  
/note="assembly\_fragment"

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## misc\_feature

## ORIGIN

## Alignment Scores:

Pred. No.: 3,88e-59 Length: 155869  
Score: 74.00 Matches: 74  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.18% Indels: 0  
DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x AC023420 (1-155869)

Qy 624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspLeuValPhe 643

Db 10396 AGCACCATGATGTGTCGGTGGTTTCATAGTCTCTCTCTGAGCACATCTCTGGTATTC 10337

Qy 644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGly 663

Db 10336 ACCTCGGAACAGTGTGATGCGCATCGAGTGAAGCCGCTTGTGTAGCAGCAGACCTGGT 10277

Qy 664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683

Db 10276 ATTCGGTGTGTGTGGAACACAGGTGTCTCTAGTGTATCTCTGGGGCGCTGGCACTGAT 10217



```

QY 684 GluInGluGluLysLeuLysSerGluCysPheSerLysArg 697
|||||
Db 10216 GAACAGAGAAAGTAAATCAGATGTTTTCACAAAGA 10175

RESULT 40
LOCUS AB069688S15 1207 bp DNA linear ROD 04-APR-2002
DEFINITION Rattus norvegicus Atrn gene for membrane attractin, exon 16.
ACCESSION AB049236
VERSION AB049236.1 GI:12275377
KEYWORDS
SEGMENT
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
REFERENCE
1 Rattus norvegicus
AUTHORS Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T., Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T. and Serikawa, T.
TITLE Attractin/mahogany/zitter plays a critical role in myelination of the central nervous system
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
MEDLINE 21143347
PUBMED 11209055
REFERENCE
2 (bases 1 to 1207)
AUTHORS Kuramoto, T. and Ushijima, T.
TITLE Direct Submission
JOURNAL Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center Research Institute, Carcinogenesis Division, Teukiji 5-1-1, Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@gan2.ncc.go.jp, Tel:81-33542-2511, Fax:81-5565-1753)
FEATURES
source
1..1207
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/chromosome="3"
488..707
/gene="Atrn"
/product="membrane attractin"
/note="CDS is reported in Acc# AB049248"
/number=16
exon

ORIGIN
Alignment Scores: 7,24e-60 Length: 1207
Pred. No.: 73.00 Matches: 73
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 5.11% Gaps: 0
DB: 10

US-09-787-097-12 (1-1429) x AB069688S15 (1-1207)

QY 849 SerLysLeuThrLeuThrProTTPValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
|||||
Db 488 TCCAGCTCAGCTGACTGCTCATGGTGGTCTTCGGAAGATCAATGATCTTACTGCTGC 547

QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
|||||
Db 548 TGGGAGGATPATCTCTCCATTCACAATAGTTTGTGTCAGTGGATGCCATCTGAGCCCACT 607

QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
|||||
Db 608 GATGCGGGTCTTGTGGGATCTTGTCAGAGCCTAGTACTCGGGGACTGAAGCTGCAACC 667

QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
|||||
Db 668 TGCATCAACCCCACTCAATGGAAGTGTGTGTGAAGGCCT 706

RESULT 41
LOCUS H009384S14 1728 bp DNA linear PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 16.
ACCESSION AF218902
VERSION AF218902.1 GI:8118068
KEYWORDS
SEGMENT
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
1 (bases 1 to 1728)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Secreted and membrane attractin result from alternative splicing of the human ATRN gene
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
MEDLINE 20283944
PUBMED 10811918
REFERENCE
2 (bases 1 to 1728)
AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
TITLE Direct Submission
JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
FEATURES
Location/Qualifiers
1..1728
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/map="20p13"
533..752
/gene="ATRN"
/number=16
exon

ORIGIN
Alignment Scores: 9,65e-60 Length: 1728
Pred. No.: 73.00 Matches: 73
Score: 100.00% Conservative: 0
Percent Similarity: 100.00% Mismatches: 0
Best Local Similarity: 100.00% Indels: 0
Query Match: 5.11% Gaps: 0
DB: 9

US-09-787-097-12 (1-1429) x H009384S14 (1-1728)

QY 849 SerLysLeuThrLeuThrProTTPValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
|||||
Db 533 TCCAGCTCAGCTTACCCCACTGGTGGCTTCGGAAGATCAATGATCTTACTGCTGC 592

QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
|||||
Db 593 TGGGAAGATATGCCCCATTTACAAATAGTTTACTACGTGGATGCCGTCTGAGCCCACT 652

QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
|||||
Db 653 GATGCTGATCTTGTGGAATTTATTCAGAACCCCACTACTCGGGGACTGAAGGCTGCAACC 712

QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
|||||
Db 713 TGCATCAACCCCACTCAATGTTAGTGTGTGTAAGGCCT 751

RESULT 42
LOCUS AR164810 9050 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 5 from patent US 6274339.
ACCESSION AR164810
VERSION AR164810.1 GI:16238016
KEYWORDS
SOURCE Unknown.

```

30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) RP23-228H14 is from the RPI-23 Mouse PAC Library<sup>2</sup> constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6.

FEATURES  
source  
1..90050  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN

Alignment Scores:  
Pred. No.: 2,29e-58 Length: 19283  
Score: 73.00 Matches: 73  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.11% Indels: 0  
DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x AR164810 (1-90050)

QY 849 SerLysLeuThrLeuThrProTrrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
Db 28996 TCCAAGCTCCTGACTCCATGGTGTCTTCGGAAGATCAATGTGCTTACTGGTGC 29055  
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
Db 29056 TGGAGGATATGTCCTCATCAATAGTTGCTGCAATGTCATGTCAGCCAGT 29115  
QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaIaThr 908  
Db 29116 GATCGTGGTCTTCTGTGGATCTTGTTCAGAGCCTAGTACTCGGGGATTAAGGCTGCAACC 29175  
QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
Db 29176 TGCAATCAACCTCTCAATGCGAGCGTCTGTGAAGCCCT 29214

Alignment Scores:  
Pred. No.: 2,29e-58 Length: 90050  
Score: 73.00 Matches: 73  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.11% Indels: 0  
DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x AL833771 (1-189283)

QY 849 SerLysLeuThrLeuThrProTrrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
Db 28996 TCCAAGCTCCTGACTCCATGGTGTCTTCGGAAGATCAATGTGCTTACTGGTGC 29055  
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
Db 29056 TGGAGGATATGTCCTCATCAATAGTTGCTGCAATGTCATGTCAGCCAGT 29115  
QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaIaThr 908  
Db 29116 GATCGTGGTCTTCTGTGGATCTTGTTCAGAGCCTAGTACTCGGGGATTAAGGCTGCAACC 29175  
QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
Db 29176 TGCAATCAACCTCTCAATGCGAGCGTCTGTGAAGCCCT 29214

US-09-787-097-12 (1-1429) x AL833771 (1-189283)

QY 849 SerLysLeuThrLeuThrProTrrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
Db 28996 TCCAAGCTCCTGACTCCATGGTGTCTTCGGAAGATCAATGTGCTTACTGGTGC 29055  
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
Db 29056 TGGAGGATATGTCCTCATCAATAGTTGCTGCAATGTCATGTCAGCCAGT 29115  
QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaIaThr 908  
Db 29116 GATCGTGGTCTTCTGTGGATCTTGTTCAGAGCCTAGTACTCGGGGATTAAGGCTGCAACC 29175  
QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
Db 29176 TGCAATCAACCTCTCAATGCGAGCGTCTGTGAAGCCCT 29214

US-09-787-097-12 (1-1429) x AL833771 (1-189283)

QY 849 SerLysLeuThrLeuThrProTrrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
Db 28996 TCCAAGCTCCTGACTCCATGGTGTCTTCGGAAGATCAATGTGCTTACTGGTGC 29055  
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
Db 29056 TGGAGGATATGTCCTCATCAATAGTTGCTGCAATGTCATGTCAGCCAGT 29115  
QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaIaThr 908  
Db 29116 GATCGTGGTCTTCTGTGGATCTTGTTCAGAGCCTAGTACTCGGGGATTAAGGCTGCAACC 29175  
QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
Db 29176 TGCAATCAACCTCTCAATGCGAGCGTCTGTGAAGCCCT 29214

RESULT 43  
AL833771  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AL833771 189283 bp DNA linear ROD 20-AUG-2002  
Mouse DNA sequence from clone RP23-228H14 on chromosome 2, complete sequence.

AL833771 AL833771.4 GI:22449827  
HTG.  
Mus musculus (house mouse)

Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 189283)

Direct Submission  
Submitted (20-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
On Aug 22, 2002 this sequence version replaced gi:22416184.

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: humquerry@sanger.ac.uk

RESULT 44  
AC094932/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

AC094932 245589 bp DNA linear HTG 09-MAY-2003  
Rattus norvegicus clone CH230-6D19, \*\*\* SEQUENCING IN PROGRESS \*\*\*.

AC094932 AC094932 GI:30466943  
HTG; HTGS PHASE2; HTGS DRAFT; HTGS ENRICHED.  
Rattus norvegicus (Norway rat)

Rattus norvegicus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.  
1 (bases 1 to 245589)

Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Altschrocks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasat, H., Chen, Y., Chen, Z.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

REFERENCE  
AUTHORS

1 (bases 1 to 245589)  
Muzny, D. Marie., Metzker, M. Lee., Abramson, S., Adams, C., Alder, J.,  
Allen, C., Allen, H., Altschrocks, S., Amin, A., Anguiano, D.,  
Anyalebechi, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F.,  
Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E.,  
Cardenas, V., Carter, K., Cavazos, I., Ceasat, H., Chen, Y., Chen, Z.,  
Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K.,

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above. This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >=

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, I., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Georegeorgis, E., Geet, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hughes, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowals, C., Kratt, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensheva, L., Louised, H., Lozado, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartine, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mahoney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milsavljivic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwankweli, O., Okwuonu, G., Olarnpunagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfankuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puazo, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Uemami, K., Valas, R., Vera, V., Villalana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, E., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausern, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

# TITLE JOURNAL REFERENCE AUTHORS TITLE JOURNAL

## REFERENCE AUTHORS TITLE JOURNAL

### COMMENT

----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: GBTI  
Center clone name: CH230-6D19

#### ----- Summary Statistics

Assembly program: Atlas;  
Consensus quality: 205514 bases at least Q40  
Consensus quality: 207203 bases at least Q30  
Consensus quality: 208050 bases at least Q20  
Estimated insert size: 221727; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbankdraft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbankdraft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.

#### FEATURES

##### source

1..245589  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-6D19"  
545..1006  
/note="clone boundary  
clone end:Sp6  
site:EcoRI  
end sequence:BH358667"  
209973..210849  
/note="clone boundary  
clone end:T7  
site:EcoRI  
end sequence:BH358665"

##### misc\_feature

##### misc\_feature

#### ORIGIN

Alignment Scores:  
Pred. No.: 5,12e-58 Length: 245589  
Score: 73.00 Matches: 73  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.11% Indels: 0  
DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x AC094932 (1-245589)

QY 849 SerlystleuthrThrProTtpValglyLeuArglyleAsnValSerTyTrpCys 868  
DB 99645 TCCAGGTCACTCTGACTCCATTCACAAATAGTTTGTGAGTGGATGCCATCTGAGCCAGT 99586  
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
DB 99585 TGGGAGGATATGCTCTCCATTCACAAATAGTTTGTGAGTGGATGCCATCTGAGCCAGT 99526  
QY 889 AspalaglyPheCysGlylleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908  
DB 99525 GATCCGGCTCTGTGGGATCTTTGTACAGACTAGTACTCGGGACTGAAGGCTGCAACC 99466  
QY 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
DB 99465 TGCATCAACCACTCATTGGAGTGTGTGTAAGGCT 99427

#### RESULT 45

H009384S08

LOCUS

H009384S08

DEFINITION

Homosapiens attractin precursor (ATRN) gene, exon 8.

ACCESSION

AF218896

VERSION

AF218896.1

GI:8118062

KEYWORDS

8 of 27

SEGMENT

SOURCE

Homosapiens (human)



transcription elongation.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 51371)  
Ellington,A  
Direct Submission  
Submitted (19-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,  
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
Requests: clonesrequest@sanger.ac.uk  
On Oct 1, 2000 this sequence version replaced gi:19930864.  
During differences assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations.  
Together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; Information  
on the WORMPEP database can be found at:  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep/](http://www.sanger.ac.uk/Projects/C_elegans/wormpep/) This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 20, constructed by the Sanger Centre Chromosome 20  
Mapping Group. Further information can be found at  
<http://www.sanger.ac.uk/HGP/Chr20>  
IMPORTANT: This sequence is not the entire insert of clone  
RP11-137F15. It may be shorter because we sequence overlapping  
sections only once, except for a 100 base overlap.  
The true left end of clone RP4-581P3 is at 51272 in this sequence.  
The true right end of clone RP4-687F11 is at 100 in this sequence.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30);  
an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. RP11-137F15 is from  
the library RPC1-11.1 constructed by the group of Pieter de Jong.  
For further details see  
<http://www.chori.org/bacpac/home.htm>  
VECTOR: pBACe3.6.  
Location/Qualifiers  
1..51371  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/clone="RP11-137F15"  
/clone\_lib="RPC1-11.1"  
11..374  
/note="LIPB3 repeat: matches 5787..6143 of consensus"  
repeat\_region  
483..562  
/note="40 copies 2 mer ta 63% conserved"  
repeat\_region  
489..660  
/note="43 copies 4 mer tata 57% conserved"  
repeat\_region  
620..661  
/note="21 copies 2 mer at 81% conserved"  
misc\_feature  
3375..3794  
/note="match: GSS: Em:AQ755239"  
repeat\_region  
4861..5206  
/note="MER41D repeat: matches 169..548 of consensus"  
repeat\_region  
5660..5699  
/note="MER41D repeat: matches 101..140 of consensus"  
repeat\_region  
5943..6017  
/note="MER41D repeat: matches 1..75 of consensus"  
repeat\_region  
15032..15083  
/note="LTR37B repeat: matches 233..284 of consensus"  
repeat\_region  
15142..15189  
/note="12 copies 4 mer acac 97% conserved"  
repeat\_region  
15489..15537  
/note="MER41D repeat: matches 215..263 of consensus"

```
/gene="ATRN"
/notes="continues in En:AL109805 as dJ581P3.2
match: proteins: Tr:O75882 Tr:O95414 Tr:O9WU77 Tr:O9WU60"
/codon_start=1
/evidence=not experimental
/product="bAl37F15.2 (attractin (with dipeptidylpeptidase
IV activity))"
/protein_id="CAC17563.1"
/db_xref="GI:11493257"
/translation="MVNAAATAEALRRRTAATAALAGSGPFWHDVTRAGPGLG
AGRLPRLSPRLRLILLLLPPLLLPCEAAAAAASVSSAAAEKCD
RCVNGVGRNPGTCQCVAGWVGQCQHCGRF"
48760.48849
/notes="30 copies 3 mer gct 75% conserved"
repeat_region
48862..48903
/notes="16 copies 3 mer ggc 77% conserved"
ORIGIN
Alignment Scores:
Pred. No.: 9 52e-54 Length: 51371
Score: 68.00 Matches: 136
Percent Similarity: 98.55% Conservative: 0
Best Local Similarity: 98.55% Mismatches: 1
Query Match: 4.76% Indels: 2
DB: 9 Gaps: 0
US-09-787-097-12 (1-1429) x AL353193 (1-51371)
QY 1 MetValAlaAlaAlaAlaThrGluAlaArgLeuArgArgArgThrAlaAlaThrAla 20
Db 48617 ATGCTGGCTCGAGCGCGCAATGAGCAAGCGTGGAGAGAGAGCGCGCGACGCA 48676
QY 21 AlaLeuAlaGlyArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg 40
Db 48677 GCCTCGCGGCGAGGAGCGCGCGCGCTGCGACTGGGACGTGACCAAGGCTGGGAGG 48736
QY 41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60
Db 48737 CCGGGGCTGGGGCGCGGCTGGCGCTCCCGCGGCTGCTCTCCACCGCTGGCGGACCGG 48796
QY 61 LeuLeuLeuLeuLeuLeuLeuPro-ProProLeuLeuLeuLeuLeuProCysGly 80
Db 48797 CTGCTGCTGCTGCTGTGTGTGTCTTC-GCGCGCGCTGCTGCTGCTGCTGCTGCTG 48855
QY 80 uAlaGluAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGly 100
Db 48856 GGCCGAGGCGCGCGCGCGCGCGGTGCTGCGGCTCAGCGGACGCGGCGCAAGGA 48915
QY 100 uCysAspArgProCysValAsnGlyCysValAsnProGlyThrGlyGlnCysValC 120
Db 48916 ATGTGACCGGCGCTGTGTCAACGGCGGTGCTGCAACCTGGCACCCTGGCTGCTG 48975
QY 120 sProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArg 137
Db 48976 CCCCCTGGCTGGTGGCGGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGG 49027
RESULT 48
AC074061/c 63513 bp DNA linear HTG 13-JUL-2000
LOCUS Homo sapiens chromosome 2 clone RP11-770B8 map 2, LOW-PASS SEQUENCE
DEFINITION
AC074061
VERSION AC074061.1 GI:9150468
KEYWORDS HTG; HTGS_PHASE0.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 63513)
REFERENCE
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 2, clone RP11-770B8
JOURNAL Unpublished
```

REFERENCE  
AUTHORS

2 (bases 1 to 63513)  
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Bastien,V., Beda,F.,  
Boguski,M., Bouckert,B., Brown,A., Burkett,G.,  
Campione,A., Castle,A., Choepe,I., Colangelo,M., Collins,S.,  
Coddie,S., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., FitzHugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,  
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lechazky,J.,  
Levine,R., Liu,G., Liu,G., Locke,K., MacDonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,  
Meldrim,J., Meneus,L., Mihova,T., Miranda,C., Mienga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Tesfaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (13-JUL-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RN/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIER  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L5682  
Center clone name: 770\_E8

TITLE  
JOURNAL  
COMMENT

\* NOTE: This record contains 72 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1 780: contig of 780 bp in length  
\* 781 880: gap of 100 bp  
\* 881 1658: contig of 778 bp in length  
\* 1659 1758: gap of 100 bp  
\* 1759 2556: contig of 798 bp in length  
\* 2557 2656: gap of 100 bp  
\* 2657 3437: contig of 781 bp in length  
\* 3438 3537: gap of 100 bp  
\* 3538 4308: contig of 771 bp in length  
\* 4309 4408: gap of 100 bp  
\* 4409 5165: contig of 757 bp in length  
\* 5166 5265: gap of 100 bp  
\* 5266 6031: contig of 766 bp in length  
\* 6032 6131: gap of 100 bp  
\* 6132 6896: contig of 765 bp in length  
\* 6897 6996: gap of 100 bp  
\* 6997 7771: contig of 775 bp in length  
\* 7772 7871: gap of 100 bp  
\* 7872 8651: contig of 780 bp in length  
\* 8652 8751: gap of 100 bp  
\* 8752 9524: contig of 773 bp in length  
\* 9525 9624: gap of 100 bp  
\* 9625 10447: contig of 823 bp in length  
\* 10448 10547: gap of 100 bp  
\* 10548 11406: contig of 859 bp in length





```

AC023420.4  GI:10280876
VERSION      HTGS PHASE1; HTGS_DRAFT.
KEYWORDS     Homo sapiens (human)
SOURCE       Homo sapiens
ORGANISM     Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1 (bases 1 to 155869)
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE        Homo sapiens chromosome 2, clone RP11-737G10
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 155869)
AUTHORS      Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
              Anderson,S., Baldwin,J., Barna,N., Bedalov,A., Boguslavsky,L.,
              Boukhgalter,B., Brown,A., Burkett,G., Campopiano,A., Castle,A.,
              Choquel,Y., Collinge,M., Collins,S., Collins,A., Cooke,P.,
              DeArrellano,K., Dewar,K., Dodge,S., Domino,M., Doyle,M.,
              Fenestor,J., Ferreira,P., FitzHugh,W., Forrest,C., Gage,D.,
              Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,
              Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
              Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,
              Klein,J., Landers,T., Largocque,K., Lehoczy,J., Levine,R.,
              Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N., McCarthy,M.,
              McEwan,P., McGurk,A., McKernan,K., McPheeters,R., Meldrum,J.,
              Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J., Naylor,J.,
              Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Olivari,T.M.,
              Peterson,K., Pierre,N., Pisani,C., Pollara,V., Raymond,C.,
              Riley,R., Rogov,P., Rothman,D., Roy,A., Santos,R., Schauer,S.,
              Severly,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
              Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,
              Travers,M., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B.,
              Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zimmer,A. and
              Zody,M.
DIRECT SUBMISSION
Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 23, 2000 this sequence version replaced gi:7767820.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L6805
Center clone name: 737_G_10
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 145736 bases at least Q40
Consensus quality: 151534 bases at least Q30
Consensus quality: 153551 bases at least Q20
Insert size: 157000; agarose-ff
Quality coverage: 4.2 in Q20 bases; agarose-ff
Quality coverage: 4.3 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 12284: contig of 12284 bp in length
* 12285: gap of 100 bp
* 12385: contig of 1320 bp in length
* 13705: gap of 100 bp
* 13805: contig of 1853 bp in length
* 15658: gap of 100 bp

```

```

17755: contig of 1998 bp in length
17855: gap of 100 bp
19875: contig of 2020 bp in length
19975: gap of 100 bp
22894: contig of 2919 bp in length
22994: gap of 100 bp
26992: contig of 3998 bp in length
27092: gap of 100 bp
31422: contig of 4330 bp in length
31522: gap of 100 bp
37920: contig of 6398 bp in length
38020: gap of 100 bp
48644: contig of 10624 bp in length
48744: gap of 100 bp
69112: contig of 20368 bp in length
69212: gap of 100 bp
86049: contig of 16837 bp in length
86149: gap of 100 bp
107477: contig of 21328 bp in length
107577: gap of 100 bp
131094: contig of 23517 bp in length
131094: gap of 100 bp
155869: contig of 24675 bp in length.
131195 Location/Qualifiers
          1..155869
             /organism="Homo sapiens"
             /mol_type="genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="2"
             /map="2"
             /clone="RP11-737G10"
             /clone_lib="RP11-11 Human Male BAC"
          1..12284
             /note="assembly_fragment"
             clone_end:SP6
             vector_side:left
          12385..13704
             /note="assembly_fragment"
          13805..15657
             /note="assembly_fragment"
          15758..17755
             /note="assembly_fragment"
          17856..19875
             /note="assembly_fragment"
          19976..22894
             /note="assembly_fragment"
          22995..26992
             /note="assembly_fragment"
          27093..31422
             /note="assembly_fragment"
          31523..37920
             /note="assembly_fragment"
          38021..48644
             /note="assembly_fragment"
          48745..69112
             /note="assembly_fragment"
          69213..86049
             /note="assembly_fragment"
          86150..107477
             /note="assembly_fragment"
          107578..131094
             /note="assembly_fragment"
          131195..155869
             /note="assembly_fragment"
ORIGIN
Alignment Scores:
Pred. No.:      2.32e-53      Length:      155869
Score:          68.00        Matches:      136
Percent Similarity: 98.55%    Conservative: 0
Best Local Similarity: 98.55%  Mismatches:    1
Query Match:      4.76%      Indels:       2
DB:              2          Gaps:        0

```



\* NOTE: This record contains 125 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.

1  
856: contig of 856 bp in length  
857 956: gap of 100 bp  
957 1821: contig of 865 bp in length  
1822 1921: gap of 100 bp  
1922 2776: contig of 855 bp in length  
2777 2876: gap of 100 bp  
2877 3767: contig of 891 bp in length  
3768 3867: gap of 100 bp  
3868 4709: contig of 842 bp in length  
4710 4809: gap of 100 bp  
4810 5655: contig of 845 bp in length  
5656 5755: gap of 100 bp  
5756 6627: contig of 872 bp in length  
6628 6727: gap of 100 bp  
6728 7586: contig of 859 bp in length  
7587 7687: gap of 100 bp  
7688 8541: contig of 855 bp in length  
8542 8641: gap of 100 bp  
8642 9524: contig of 883 bp in length  
9525 9624: gap of 100 bp  
9625 10509: contig of 885 bp in length  
10510 10609: gap of 100 bp  
10610 11474: contig of 865 bp in length  
11475 11574: gap of 100 bp  
11575 12441: contig of 867 bp in length  
12442 12541: gap of 100 bp  
12542 13409: contig of 868 bp in length  
13410 13509: gap of 100 bp  
13510 14439: contig of 930 bp in length  
14440 14539: gap of 100 bp  
14540 15433: contig of 894 bp in length  
15434 15533: gap of 100 bp  
15534 16442: contig of 909 bp in length  
16443 16542: gap of 100 bp  
16543 17421: contig of 879 bp in length  
17422 17521: gap of 100 bp  
17522 18380: contig of 859 bp in length  
18381 18480: gap of 100 bp  
18481 19345: contig of 865 bp in length  
19346 19445: gap of 100 bp  
19446 20327: contig of 882 bp in length  
20328 20427: gap of 100 bp  
20428 21237: contig of 870 bp in length  
21298 21397: gap of 100 bp  
21398 22055: contig of 858 bp in length  
22056 22155: gap of 100 bp  
22156 23005: contig of 850 bp in length  
23006 23105: gap of 100 bp  
23106 24016: contig of 911 bp in length  
24017 24116: gap of 100 bp  
24117 24972: contig of 856 bp in length  
24973 25072: gap of 100 bp  
25072 25977: contig of 905 bp in length  
25978 26077: gap of 100 bp  
26078 26952: contig of 875 bp in length  
26953 27052: gap of 100 bp  
27053 27911: contig of 859 bp in length  
27912 28011: gap of 100 bp  
28012 28668: contig of 857 bp in length  
28669 28968: gap of 100 bp  
28969 29624: contig of 656 bp in length  
29625 29724: gap of 100 bp

29725 30574: contig of 850 bp in length  
30575 30674: gap of 100 bp  
30675 31530: contig of 856 bp in length  
31531 31630: gap of 100 bp  
31631 32491: contig of 861 bp in length  
32492 32591: gap of 100 bp  
32592 33467: contig of 876 bp in length  
33468 33567: gap of 100 bp  
33568 34437: contig of 870 bp in length  
34438 34537: gap of 100 bp  
34538 35404: contig of 867 bp in length  
35405 35504: gap of 100 bp  
35505 36368: contig of 864 bp in length  
36369 36468: gap of 100 bp  
36469 37353: contig of 885 bp in length  
37354 37453: gap of 100 bp  
37454 38310: contig of 857 bp in length  
38311 38410: gap of 100 bp  
38411 39267: contig of 857 bp in length  
39268 39367: gap of 100 bp  
39368 40217: contig of 850 bp in length  
40218 40317: gap of 100 bp  
40318 41226: contig of 909 bp in length  
41227 41326: gap of 100 bp  
41328 42195: contig of 869 bp in length  
42196 42295: gap of 100 bp  
42296 43180: contig of 885 bp in length  
43181 43280: gap of 100 bp  
43281 44143: contig of 863 bp in length  
44144 44243: gap of 100 bp  
44244 45131: contig of 888 bp in length  
45132 45231: gap of 100 bp  
45232 46149: contig of 918 bp in length  
46150 46249: gap of 100 bp  
46250 47122: contig of 873 bp in length  
47123 47222: gap of 100 bp  
47223 48101: contig of 879 bp in length  
48102 48201: gap of 100 bp  
48202 49041: contig of 840 bp in length  
49042 49141: gap of 100 bp  
49142 50059: contig of 918 bp in length  
50060 50159: gap of 100 bp  
50160 51029: contig of 870 bp in length  
51030 51129: gap of 100 bp  
51130 51995: contig of 866 bp in length  
51996 52095: gap of 100 bp  
52096 52971: contig of 876 bp in length  
52972 53071: gap of 100 bp  
53072 53942: contig of 871 bp in length  
53943 54042: gap of 100 bp  
54043 54923: contig of 881 bp in length  
54924 55023: gap of 100 bp  
55024 55877: contig of 854 bp in length  
55878 55977: gap of 100 bp  
55979 56851: contig of 874 bp in length  
56852 56951: gap of 100 bp  
56952 57808: contig of 857 bp in length  
57809 57908: gap of 100 bp  
57909 58755: contig of 847 bp in length  
58756 58855: gap of 100 bp  
58856 59709: contig of 854 bp in length  
59710 59809: gap of 100 bp  
59810 60521: contig of 712 bp in length  
60522 61501: gap of 100 bp  
61502 61601: gap of 100 bp  
61602 62490: contig of 889 bp in length  
62491 62590: gap of 100 bp  
62591 63450: contig of 860 bp in length  
63451 63550: gap of 100 bp  
63551 64094: contig of 544 bp in length  
64095 64194: gap of 100 bp  
64195 65022: contig of 828 bp in length



Percent Similarity: 100.00%	Conservative: 0
Best Local Similarity: 100.00%	Mismatches: 0
Query Match: 3.99%	Indels: 0
DB: 9	Gaps: 0
US-09-787-097-12 (1-1429) x H009384506 (1-1032)	
QY 315 GlyProGlyCySerValProValProAlaAsnGlnSerPheThrTrpThrArgGluGluTyr 334	
Db 528 GGTCTCTGGATGTTTCAGTTCCTCTTACACGCTAACCAAGTCATTTTGGACTCGAGAGGAATAT 587	
QY 335 SerAsnLeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrp 354	
Db 588 TCTACTTAAAGTCCCGAGACATCTATAAAGCTGGTCATGTGGAACATTTATGTGG 647	
QY 355 ValValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAla 371	
Db 648 GTTGTGGAGGATATATGTTCAACCACTCAGATTATAACATGGTTCAGCG 698	
RESULT 54	
H009384521	
LOCUS H009384521	
DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 24.	
ACCESSION AF218909	
VERSION AF218909.1	GI:8118075
KEYWORDS	
SEGMENT	
SOURCE	21 of 27
ORGANISM	Homo sapiens (human)
REFERENCE	1 (bases 1 to 1976)
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.
TITLE	Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
JOURNAL	Secreted and membrane attractin result from alternative splicing of the human ATRN gene
MEDLINE	Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)
PUBMED	20283944
REFERENCE	10811918
AUTHORS	2 (bases 1 to 1976)
TITLE	Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.
JOURNAL	Direct Submission
FEATURES	Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA
source	Location/Qualifiers
exon	1..1976
	/organism="Homo sapiens"
	/mol_type="genomic DNA"
	/db_xref="taxon:9606"
	/chromosome="20"
	/map="20p13"
	/feature="ATRN"
	/number=24
exon	392..549
	/feature="ATRN"
	/number=24
ORIGIN	
Alignment Scores:	
Pred. No.:	1.93e-40
Score:	53.00
Percent Similarity:	100.00%
Best Local Similarity:	100.00%
Query Match:	3.71%
DB:	9
US-09-787-097-12 (1-1429) x H009384521 (1-1976)	
QY 1215 AlaGlyThrGlnAlaGlyGluMetProValValSerIleThrAsnIleLysGluTyr 1234	
Db 391 GCTGAACCCAGGCTGGAGAGAGATGCTCTGTTTCAAAACCAACATTAAGGAGTAC 450	
QY 1235 LysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePhe 1254	

Db 3035 AAAGATAGCTTCTCAATGAGAAATTTGATTTTCGCAACCACTTCTTCTT 3094  
QY 1255 ValTyValSeAsnPhetTrpPrrolleLysIleGln 1267  
Db 3095 GTTATGTCAGTAATTTACCTGGCCCATCAAAATTCAG 3133

RESULT 56  
AC055771  
DEFINITION Homo sapiens chromosome 20 clone RP11-561P16 map 20, WORKING DRAFT  
AC055771  
ACCESSION AC055771.2 GI:7770449  
VERSION HTG; HTGS PHASE1; HTGS\_DRAFT.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 167357)  
Birken, B., Linton, L., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 20, clone RP11-561P16  
Unpublished  
2 (bases 1 to 167357)  
Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., Mienga, V., Morrow, J.,  
Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (18-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 167357)  
Birken, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,  
Boguslavskiy, L., Boukhgalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., DeArellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., FitzHugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRoque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Liu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., Mienga, V., Morrow, J.,  
Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Tessaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA

## COMMENT

On May 12, 2000 this sequence version replaced gi:7582593.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center Project name: L5007  
Center clone name: 561\_P\_16  
----- Summary Statistics  
Sequencing vector: M13; M7815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 156504 bases at least Q40  
Consensus quality: 161679 bases at least Q30  
Consensus quality: 163964 bases at least Q20  
Insert size: 170000; agarose-fp  
Quality coverage: 165357; sum-of-contigs  
Quality coverage: 4.5 in Q20 bases; agarose-fp  
Quality coverage: 4.6 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 21 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

\* 1 1216: contig of 1216 bp in length  
\* 1217 1316: gap of 100 bp  
\* 1317 2985: contig of 1669 bp in length  
\* 2986 3085: gap of 100 bp  
\* 3086 4628: contig of 1543 bp in length  
\* 4629 4728: gap of 100 bp  
\* 4729 7507: contig of 2779 bp in length  
\* 7508 7607: gap of 100 bp  
\* 7608 11292: contig of 3685 bp in length  
\* 11293 11392: gap of 100 bp  
\* 11393 14213: contig of 2821 bp in length  
\* 14214 14313: gap of 100 bp  
\* 14314 17828: contig of 3515 bp in length  
\* 17829 17928: gap of 100 bp  
\* 17929 21068: contig of 3140 bp in length  
\* 21069 21168: gap of 100 bp  
\* 21169 25035: contig of 3867 bp in length  
\* 25036 25135: gap of 100 bp  
\* 25136 28331: contig of 3196 bp in length  
\* 28332 28431: gap of 100 bp  
\* 28432 33595: contig of 5164 bp in length  
\* 33596 33695: gap of 100 bp  
\* 33696 40346: contig of 6651 bp in length  
\* 40347 40446: gap of 100 bp  
\* 40447 43807: contig of 3361 bp in length  
\* 43808 43907: gap of 100 bp  
\* 43908 48727: contig of 4820 bp in length  
\* 48728 48827: gap of 100 bp  
\* 48828 56083: contig of 7256 bp in length  
\* 56084 56183: gap of 100 bp  
\* 56184 67065: contig of 10882 bp in length  
\* 67066 67165: gap of 100 bp  
\* 67166 82881: contig of 15716 bp in length  
\* 82882 82981: gap of 100 bp  
\* 82982 101845: contig of 18864 bp in length  
\* 101846 101945: gap of 100 bp  
\* 101946 122755: contig of 20810 bp in length  
\* 122756 122855: gap of 100 bp  
\* 122856 142457: contig of 19602 bp in length  
\* 142458 142557: gap of 100 bp  
\* 142558 167357: contig of 24800 bp in length.

FEATURES  
source

Location/Qualifiers  
 1. .167357  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20"  
 /clone="RP11-561P16"  
 /clone\_lib="RPC1-11 Human Male BAC"  
 1. .1216  
 /notes="assembly\_fragment"  
 1317. .2985  
 /note="assembly\_fragment"  
 3086. .4828  
 /note="assembly\_fragment"  
 4729. .7507  
 /note="assembly\_fragment"  
 7608. .11292  
 /note="assembly\_fragment"  
 11393. .14213  
 /note="assembly\_fragment"  
 14314. .17828  
 /note="assembly\_fragment"  
 17929. .21068  
 /note="assembly\_fragment"  
 21169. .25035  
 /note="assembly\_fragment"  
 25136. .28331  
 /note="assembly\_fragment"  
 28432. .33595  
 /note="assembly\_fragment"  
 33696. .40346  
 /note="assembly\_fragment"  
 40447. .43807  
 /note="assembly\_fragment"  
 clone\_end:SP6  
 vector\_side:right  
 43908. .48727  
 /note="assembly\_fragment"  
 48828. .56083  
 /note="assembly\_fragment"  
 56184. .67065  
 /note="assembly\_fragment"  
 67166. .82881  
 /note="assembly\_fragment"  
 82982. .101845  
 /note="assembly\_fragment"  
 clone\_end:T7  
 vector\_side:right  
 101946. .122755  
 /note="assembly\_fragment"  
 122856. .142457  
 /note="assembly\_fragment"  
 142558. .167357  
 /note="assembly\_fragment"

## ORIGIN

Alignment Scores:  
 Pred. No.: 6,77e-39 Length: 167357  
 Score: 53.00 Matches: 53  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.71% Indels: 0  
 DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x AC055771 (1-167357)

QY 1215 AlaGlyThrGlnAlaGlyGluGluMetProValValSerIleThrAsnIleLysGluTyr 1234  
 Db 2354 GCTGGAAACCGAGCTGGAGAGATGCTGTGTTTCAAAACCAACATTAGAGGATC 2413  
 QY 1235 LysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePhe 1254

Db 2414 AAAGATAGTTTCTCTAATGAGAGTTTGATTTTCGCAACCAACCAATATACACTTTCTTT 2473  
 QY 1255 ValTyValSerAsnPhThrTTPProIleLysIleGln 1267  
 Db 2474 GTTATGTCAGTAATTTTACCTGCGCCATCAAAATTCAG 2512  
 RESULT 57  
 AC017113 206329 bp DNA linear HTG 07-JUL-2000  
 LOCUS Homo sapiens chromosome 20 clone RP11-574H7, WORKING DRAFT  
 DEFINITION SEQUENCE, 31 unordered pieces.  
 ACCESSION AC017113  
 VERSION AC017113.4 GI:9570339  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 Waterston,R.H.  
 The sequence of Homo sapiens clone  
 Unpublished  
 2 (bases 1 to 206329)  
 Waterston,R.H.  
 Direct Submission  
 Submitted (09-DEC-1999) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 ON Jun 17, 2000 this sequence version replaced gi:6855220.  
 COMMENT  
 ----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site:http://genome.wustl.edu/gsc/index.shtml  
 ----- Project Information -----  
 Center project name: H NH0574H07  
 ----- Summary Statistics -----  
 Sequencing vector: M13; 91%  
 Chemistry: Dye-primer ET; 91% of reads  
 Chemistry: Dye-terminator Big Dye; 9% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 184921 bases at least Q40  
 Consensus quality: 192397 bases at least Q30  
 Consensus quality: 195862 bases at least Q20  
 Insert size: 237000; agarose-fp  
 Insert coverage: 3.84 in Q20 bases; agarose-fp  
 Quality coverage: 3.66 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 31 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* \* 1 1539: contig of 1539 bp in length  
 \* \* 1540 1639: gap of unknown length  
 \* \* 1640 3259: contig of 1620 bp in length  
 \* \* 3260 3359: gap of unknown length  
 \* \* 3360 5400: contig of 2041 bp in length  
 \* \* 5401 5501: gap of unknown length  
 \* \* 5501 7783: contig of 2283 bp in length  
 \* \* 7784 11241: contig of 3358 bp in length  
 \* \* 11242 11341: gap of unknown length  
 \* \* 11342 13482: contig of 2141 bp in length  
 \* \* 13483 13582: gap of unknown length  
 \* \* 13583 16915: contig of 3333 bp in length  
 \* \* 16916 17015: gap of unknown length  
 \* \* 17016 21346: contig of 4331 bp in length



```
* 21347 21446: gap of unknown length
* 21447 25103: contig of 3657 bp in length
* 25104 25203: gap of unknown length
* 28415 28416: contig of 3212 bp in length
* 28516 32873: contig of 4358 bp in length
* 32874 32973: gap of unknown length
* 32974 36496: contig of 3523 bp in length
* 36497 36596: gap of unknown length
* 36597 40649: contig of 4053 bp in length
* 40650 44878: contig of 4129 bp in length
* 44879 50353: gap of unknown length
* 50354 50454: gap of unknown length
* 50455 55947: gap of unknown length
* 55948 61420: contig of 5374 bp in length
* 61421 61520: gap of unknown length
* 61521 67819: contig of 6299 bp in length
* 67820 67919: gap of unknown length
* 67920 72327: contig of 4408 bp in length
* 72328 72427: gap of unknown length
* 72428 77293: contig of 4866 bp in length
* 77294 82626: contig of 5233 bp in length
* 82627 82727: gap of unknown length
* 82728 90186: contig of 7460 bp in length
* 90187 90286: gap of unknown length
* 90287 96584: contig of 6298 bp in length
* 96585 106019: gap of unknown length
* 106020 106119: contig of 9335 bp in length
* 106120 113530: contig of 7411 bp in length
* 113531 113630: gap of unknown length
* 113631 125151: contig of 11521 bp in length
* 125152 125251: gap of unknown length
* 125252 137677: contig of 12426 bp in length
* 137678 151263: contig of 13486 bp in length
* 151264 151363: gap of unknown length
* 151364 165370: contig of 14007 bp in length
* 165371 165470: gap of unknown length
* 165471 184704: contig of 19234 bp in length
* 184705 184804: gap of unknown length
* 184805 206329: contig of 21525 bp in length.
FEATURES
  Location/Qualifiers
    1..206329
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="20"
      /clone="RP11-574H7"
ORIGIN
Alignment Scores:
Pred. No.: 8,01e-39 Length: 206329
Score: 53.00 Matches: 53
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.71% Indels: 0
DB: 2 Gaps: 0
US-09-787-097-12 (1-1429) x AC017113 (1-206329).
Qy 1215 AlaGlyThrGlnAlaGlyGluMetProValValSerIysThrAsnIleIysGluTyr 1234
Db 173249 GCTGGAACCCAGGCTGAGAGAGATGCGCTGTGTTTCAAAACCAACATTAAGAGTAC 173308
Qy 1235 LysAspSerPheSerAsnGluIysPheAspPheArgAsnHisProAsnIleThrPhePhe 1254
Db 173309 AAGATAGTTTCTCTAATGAGAGATTGATTTTCGCAACCAACCAATATCATCTTCTTT 173368
```

```
Qy 1255 ValTyrValSerAsnPheThrTyrProIleIysIleGln 1267
Db 173369 GTTATGTCAGTATTTCACCTGCCCCATCAAAATTCAG 173407
RESULT 58
LOCUS ARI164811/c ARI164811 226 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 6 from patent US 6274339.
ACCESSION ARI164811
VERSION ARI164811.1 GI:16238018
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE Unclassified.
  1 (bases 1 to 226)
AUTHORS Moore, K. and Nagle, D. Lynn.
TITLE Methods and compositions for the diagnosis and treatment of body
  weight disorders, including obesity
JOURNAL Patent: US 6274339-A 6 14-AUG-2001;
FEATURES
  source
    1..226
      /organism="unknown"
      /mol_type="unassigned DNA"
ORIGIN
Alignment Scores:
Pred. No.: 2,86e-39 Length: 226
Score: 51.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.57% Indels: 0
DB: 6 Gaps: 0
US-09-787-097-12 (1-1429) x ARI164811 (1-226)
Qy 871 AspMetSerProPheThrAsnSerLeuLeuGlnTyrMetProSerGluProSerAspAla 890
Db 225 CATATGCTCATTCACAAATAGTTTCTGAGTGGATGCCATCTGAGCCAGTGATGCT 166
Qy 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuIysAlaAlaThrCysIle 910
Db 165 GGGTCTGTGGATCTTGTCTAGAGCTAGTACTCGGGATTAAAGGCTGCAACCTGCATC 106
Qy 911 AsnProLeuAsnGlySerValCysGluArgPro 921
Db 105 AACCTCTCAATGGCAGCGCTGTGTAAAGGCT 73
RESULT 59
AB06968S09
LOCUS Rattus norvegicus Atrn gene for membrane attractin, exon 10.
DEFINITION
ACCESSION AB049230
VERSION AB049230.1 GI:12275371
KEYWORDS
SEGMENT 9 of 27
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
  Rattus norvegicus
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
  Rattus.
  1
  Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T.,
  Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T.
  and Serikawa, T.
  Attractin/mahogany/zitter plays a critical role in myelination of
  the central nervous system
  Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)
  MEDLINE 21143347
  PUBMED 11209055
  2 (bases 1 to 1050)
  Kuramoto, T. and Ushijima, T.
  Direct Submission
  TITLE Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
```

Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,  
Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@gan2.ncc.go.jp,  
Tel:81-33542-2511, Fax:81-5565-1753)

## FEATURES

source

1. .1050  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/chromosome="3"  
496. .650  
/gene="Atrn"  
/product="membrane attractin"  
/notes="CDS is reported in Acc# AB049248"  
/number=10

exon

## ORIGIN

Alignment Scores:  
Pred. No.: 9.8e-39 Length: 1050  
Score: 51.00 Matches: 51  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.57% Indels: 0  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x AB06968S09 (1-1050)

QY 545 ThrIleuLysaspSerArgPheAtgTyrLeuHsthAlaValIleValserGly 564  
DB 497 ACCATTCTTAAGACACGCCATTTTCCTACTTCATACACCTGTGATGAGTGA 556  
QY 565 ThrMetLeuValPheGlyGlyAsnThrHisAspThrSerMetSerHisGlyAlaLys 584  
DB 557 ACCATGCTGTGTTCGGAGGGAACACACAAATGACACTTCCATGAGCCAGTGC AAA 616  
QY 585 CysPheSerSerAspPheMetAlaTyrAspIle 595  
DB 617 TGCTTCTCTCAGACTTCATGGCTTATGACATT 649

RESULT 60

HS741H3S 553 bp DNA linear STS 24-AUG-1999  
LOCUS H.sapiens STS from genomic clone 741H3, sequence tagged site.

DEFINITION Al110053

ACCESSION Al110053.1 GI:5777509

VERSION STS.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1. (bases 1 to 553)  
Deloukas,P., O'Neill,D., Hoiden,J., Mistry,D., Huckle,E., Taylor,R.  
and Hunt,S.

Direct Submission

Submitted (24-AUG-1999) E-mail contact: humquery@sanger.ac.uk  
Marker stdJ741H3S96 (Primer A : AGGTGAAGCATGTGCATTC; Primer B :  
AAGAGCTCTACCTGCGC; amplicon size : 124 bp) is from sequence  
generated from the SP6 end of PAC 741H3. 741H3 is part of the  
bacterial clone contigs constructed by the Chromosome 20 Mapping  
Group. (http://www.sanger.ac.uk/HGP/Chr20/) 741H3 is from the  
library constructed at the Roswell Park Cancer Institute by the  
group of Pieter de Jong.

For further details see http://bacpac.med.buffalo.edu/.

FEATURES

source

1. .553  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/clone="741H3"

## ORIGIN

Alignment Scores:  
Pred. No.: 4.53e-36 Length: 553

Score: 48.00 Matches: 48  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.36% Indels: 0  
DB: 11 Gaps: 0

US-09-787-097-12 (1-1429) x HS741H3S (1-553)

QY 268 ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCys 287  
DB 4 ACTGTTGAATGTGAATGTTCTGAAAAGTGAAGCATGTGACATTCCTCACTGT 63  
QY 288 ThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCys 307  
DB 64 ACAGACAACCTGTGGTTTTCCTCATCGAGGCATCTGCAATTCAGTGAATGCAGAGATGC 123  
QY 308 SerCysPheSerAspTrpGlnGly 315  
DB 124 TCCTGCTTCTCAGACTGGCAGGT 147

RESULT 61

LOCUS H009384S05

DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 5.

ACCESSION AF218893

VERSION AF218893.1 GI:8118059

KEYWORDS 5 of 27

SEGMENT Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1. (bases 1 to 1232)

AUTHORS Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schlossman,S.F.

and Duke-Cohan,J.S.

Secreted and membrane attractin result from alternative splicing of

the human ATRN gene

Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)

JOURNAL MEDLINE

PUBMED 10811918

REFERENCE 2. (bases 1 to 1232)

AUTHORS Tang,W., Gunn,T.M., McLaughlin,D.F., Freeman,G.J., Barsh,G.S.,

Schlossman,S.F. and Duke-Cohan,J.S.

Direct Submission

TITLE Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,

Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,

USA

FEATURES Location/Qualifiers

1. .1232

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="20"

/map="20p13"

621. .826

/gene="ATRN"

/number=5

exon

## ORIGIN

Alignment Scores:

Pred. No.: 8.61e-36 Length: 1232  
Score: 48.00 Matches: 48  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.36% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x H009384S05 (1-1232)

QY 268 ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHisCys 287  
DB 685 ACTGTTGAATGTGAATGTTCTGAAAAGTGAAGCATGTGACATTCCTCACTGT 744

QY 288 ThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGlyCys 307  
 Db 745 ACAGACAACTGGTGGTTTCTCATCAGGCATCTCAATTCAGTCAAGTGTGACAGAGATGC 804

QY 308 SerCysPheSerAspTrpGlnGly 315  
 Db 805 TCTGCTTCTCAGACTGGCAGAGGT 828

RESULT 62  
 H009384S16  
 LOCUS Homo sapiens  
 DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 19.  
 ACCESSION AF218904  
 VERSION AF218904.1 GI:8118070  
 KEYWORDS  
 SEGMENT  
 SOURCE  
 ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 REFERENCE 1 (bases 1 to 1674)  
 AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.  
 and Duke-Cohan, J.S.  
 TITLE Secreted and membrane attractin result from alternative splicing of  
 the human ATRN gene  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)  
 MEDLINE 20283944  
 PUBMED 10811918  
 REFERENCE 2 (bases 1 to 1674)  
 AUTHORS Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,  
 Schlossman, S.F., and Duke-Cohan, J.S.  
 TITLE Direct Submission  
 JOURNAL Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,  
 Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,  
 USA

FEATURES  
 source  
 1..1674  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20p13"  
 810..947  
 /gene="ATRN"  
 /number=19

exon

ORIGIN  
 Alignment Scores:  
 Pred. No.: 9,278-34 Length: 1674  
 Score: 46.00 Matches: 46  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.22% Indels: 0  
 DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x H009384S16 (1-1674)

QY 1062 AlAcysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGlyLysCysGlu 1081  
 Db 809 GCTTGCAATGCAACGCCACACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAG 868

QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrcGlyAspProthr 1101  
 Db 869 AACCTGACCACAGCGCAGCACTCGAGACCTGCATATCTGGCTTCTACGGTATCCACCC 928

QY 1102 AspGlyCysGlyLysCysGln 1107  
 Db 929 AATGGAGGGAAATGTCTAG 946

RESULT 63  
 ARI64812/c  
 LOCUS Rattus norvegicus  
 DEFINITION Sequence 7 from patent US 6274339.

ACCESSION ARI64812  
 VERSION ARI64812.1 GI:16238020  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 221)  
 AUTHORS Moore, K. and Nagle, D. Lynn.  
 TITLE Methods and compositions for the diagnosis and treatment of body  
 weight disorders, including obesity  
 JOURNAL Patent: US 6274339-A 7 14-AUG-2001;  
 FEATURES Location/Qualifiers  
 source 1..221  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,68e-33 Length: 221  
 Score: 45.00 Matches: 45  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.15% Indels: 0  
 DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x ARI64812 (1-221)

QY 871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890  
 Db 220 GATATGTCCTCATTCACAAATAGTTTCTCAGTGGATGCACTGAGCCAGATGCT 161

QY 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaLaThrCysIle 910  
 Db 160 GGCCTTCGTGGATCTTGTCTCAGACCTAGTACTCGGGGATTAAGAGGTGCAACCTGCATC 101

QY 911 AsnProLeuAsnGly 915  
 Db 100 AACCTCTCTCAATGCT 86

RESULT 64  
 ARI64812/c  
 LOCUS Rattus norvegicus  
 DEFINITION Rattus norvegicus Atrn gene for membrane attractin, exon 17, 18.  
 ACCESSION AB049237  
 VERSION AB049237.1 GI:12275378  
 KEYWORDS  
 SEGMENT  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.  
 REFERENCE 1  
 AUTHORS Kuramoto, T., Kitada, K., Inui, T., Sasaki, Y., Ito, K., Hase, T.,  
 Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T.  
 and Serikawa, T.  
 TITLE Attractin/mahogany/zitter plays a critical role in myelination of  
 the central nervous system  
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)  
 MEDLINE 21143347  
 PUBMED 11209055  
 REFERENCE 2 (bases 1 to 1880)  
 AUTHORS Kuramoto, T. and Ushijima, T.  
 TITLE Direct Submission  
 JOURNAL Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center  
 Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,  
 Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@gan2.ncc.go.jp,  
 Tel:81-33542-2511, Fax:81-5565-1753)  
 FEATURES Location/Qualifiers  
 source 1..1880  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:10116"



US-09-787-097-12 (1-1429); x AB06968S22 (1-1474)

QY 1224 ProValValSerIysThrAsnIleIysGluTyrLysAspSerPheSerAsnGlnIysPhe 1243  
 Db 392 CCGTGGTTGTTTCAAAAACCAACATCAAGGATACAAAGATAGCTTCTTAATGAGAATTT 451  
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263  
 Db 452 GATTTCGCACCATCCAAACATTACTTCTTGTTTATGTGTCAGTAATTCACCTGGCCC 511

QY 1264 IleIysIleGln 1267

Db 512 ATCAAAATTCAA 523

#### RESULT 67

HSDJ581P3

#### LOCUS

Human DNA sequence from clone RP4-581P3 on chromosome 20. Contains an SP3A3 (splicing factor 3a, subunit 3, 60kD (SAP61)) pseudogene, part of the ATRN gene for attractin (with dipeptidylpeptidase IV activity), ESTs, STSS and GSSs, complete sequence.

#### ACCESSION

AL109805

#### VERSION

AL109805.14 GI:73211191

#### KEYWORDS

HTG; ATRN; attractin; SAP61; SP3A3.

#### SOURCE

Homo sapiens

#### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

#### REFERENCE

1 (bases 1 to 73666)

#### AUTHORS

Ramsay,H.

#### TITLE

Direct Submission

#### JOURNAL

Submitted (19-FEB-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk

#### COMMENT

requests: clonerequest@sanger.ac.uk  
 On Mar 23, 2000 this sequence version replaced gi:7228123.  
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em., EMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>

IMPORTANT: This sequence is not the entire insert of clone RP4-581P3 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.  
 The true left end of clone RP4-581P3 is at 1 in this sequence. The true left end of clone RP4-741P3 is at 73567 in this sequence. This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP4-581P3 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
 VECTOR: pCVPAC2.

#### FEATURES

##### source

1. .73666  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /clone="RP4-581P3"

repeat\_region  
 61..361  
 /clone\_lib="RPCI-4"  
 /note="L2 repeat: matches 1703..2052 of consensus"  
 repeat\_region  
 665..969  
 /note="AluY repeat: matches 1..306 of consensus"  
 repeat\_region  
 990..1292  
 /note="AluSg repeat: matches 1..303 of consensus"  
 repeat\_region  
 1551..1626  
 /note="MIR repeat: matches 162..247 of consensus"  
 repeat\_region  
 2105..2407  
 /note="AluJb repeat: matches 5..306 of consensus"  
 repeat\_region  
 2467..2599  
 /note="MIR repeat: matches 6..149 of consensus"  
 repeat\_region  
 2817..3115  
 /note="AluSg repeat: matches 3..300 of consensus"  
 repeat\_region  
 3233..3275  
 /note="L2 repeat: matches 2617..2659 of consensus"  
 repeat\_region  
 3310..3352  
 /note="L2 repeat: matches 2617..2659 of consensus"  
 repeat\_region  
 3387..3427  
 /note="L2 repeat: matches 2617..2657 of consensus"  
 repeat\_region  
 3457..3506  
 /note="L2 repeat: matches 2570..2659 of consensus"  
 repeat\_region  
 3534..3583  
 /note="L2 repeat: matches 2570..2659 of consensus"  
 repeat\_region  
 3618..3660  
 /note="L2 repeat: matches 2570..2659 of consensus"  
 repeat\_region  
 3688..3737  
 /note="L2 repeat: matches 2570..2659 of consensus"  
 repeat\_region  
 3765..3814  
 /note="L2 repeat: matches 2570..2659 of consensus"  
 repeat\_region  
 3845..4008  
 /note="L2 repeat: matches 2617..2750 of consensus"  
 repeat\_region  
 4452..4924  
 /note="L1PA7 repeat: matches 5664..6143 of consensus"  
 repeat\_region  
 5160..5428  
 /note="AluJb repeat: matches 1..304 of consensus"  
 repeat\_region  
 5463..5826  
 /note="L2 repeat: matches 2348..2706 of consensus"  
 repeat\_region  
 5827..6274  
 /note="MLT1C repeat: matches 1..466 of consensus"  
 repeat\_region  
 6275..6350  
 /note="L2 repeat: matches 2272..2348 of consensus"  
 misc\_feature  
 6783..7189  
 /note="match: GSS: Em:AQ585728"  
 misc\_feature  
 6785..6963  
 /note="match: GSS: Em:AQ585784"  
 repeat\_region  
 6874..6912  
 /note="MLT1A repeat: matches 323..361 of consensus"  
 repeat\_region  
 7594..7881  
 /note="AluJb repeat: matches 3..285 of consensus"  
 misc\_feature  
 complement(7666..8180)  
 /note="match: GSS: Em:AQ453150"  
 repeat\_region  
 8025..8050  
 /note="13 copies 2 mer tt 92% conserved"  
 repeat\_region  
 9591..9884  
 /note="AluJb repeat: matches 1..294 of consensus"  
 repeat\_region  
 10205..10511  
 /note="AluJb repeat: matches 1..307 of consensus"  
 misc\_feature  
 complement(10997..11232)  
 /note="match: STS: Em:G08041"  
 repeat\_region  
 11393..11577  
 /note="L1M4 repeat: matches 4132..4326 of consensus"  
 repeat\_region  
 11683..11848  
 /note="L1 repeat: matches 3736..3898 of consensus"  
 repeat\_region  
 11849..12157  
 /note="AluSg repeat: matches 1..309 of consensus"  
 repeat\_region  
 12158..12182  
 /note="L1 repeat: matches 3712..3736 of consensus"  
 repeat\_region  
 12200..12396  
 /note="L1M2C repeat: matches 1702..1912 of consensus"  
 repeat\_region  
 12417..12586  
 /note="L1M2S repeat: matches 7758..7913 of consensus"

```

repeat_region 12634..12914
/note="AluX repeat: matches 6..312 of consensus"
repeat_region 12958..13404
/note="L1MEC repeat: matches 1267..1711 of consensus"
repeat_region 13486..13733
/note="AluSg1 repeat: matches 36..285 of consensus"
repeat_region 14313..14415
/note="MER86 repeat: matches 89..196 of consensus"
repeat_region 14493..15186
/note="L1 repeat: matches 4409..5090 of consensus"
repeat_region 15187..15487
/note="AluX repeat: matches 12..312 of consensus"
repeat_region 15488..15539
/note="L1 repeat: matches 4358..4409 of consensus"
misc_feature 15537..15942
/note="match: STS: Em:R15876"
misc_feature 15537..15748
/note="match: STS: Em:G15134"
gene complement(15767..17243)
/gene="dJ581P3.1"
/pseudo
complement(15767..17243)
/gene="dJ581P3.1"
/note="dJ581P3.1" (SF3A3 (splicing factor 3a, subunit 3,
60kD (SAP61)) pseudogene)
match: CDNA8: Em:X81789 Em:U08815
match: EST8: Em:AI128669 Em:T56818 Em:AW470436 Em:AA300515
Em:AA703023 Em:AA760210 Em:AA054241 Em:AA443032 Em:R17811
Em:AW493824 Em:H67476 Em:AI020542 Em:AA343902 Em:H99818
Em:AA308239 Em:AA311786 Em:AI541046 Em:AA239392
Em:AA694306 Em:AA086662 Em:R15877 Em:AI226265 Em:F22705
Em:AI155866 Em:CO6258 Em:AA086076 Em:AA081903 Em:AI019200
Em:AV138764 Em:AW406960 Em:AI202796 Em:R00790 Em:AA308500
Em:AA003841 Em:AV290458 Em:AV166565 Em:AA074850
Em:AI353575 Em:AA355911 Em:AA730707 Em:AA068453 Em:T27366
Em:AA3306883 Em:AL035731 Em:AA992803 Em:AL035732
Em:AA854771 Em:AW320502 Em:AA592509 Em:AI374964
Em:AI338280 Em:AA662262 Em:AA642724 Em:W83741 Em:AA662366
Em:AA63475 Em:W35845 Em:AI269312 Em:AL118208 Em:AA986711
Em:AI374716 Em:AI908435 Em:C82711 Em:AL039268 Em:AW246145
match: proteins: Tr:Q12874 Wp:TI135 Tr:O59706 Tr:Q22469
Tr:O46106 Sw:P19736"
/pseudo
/codon_start=1
/evidence=not_experimental
misc_feature 16008..16100 GSS: Em:AQ279830"
/note="match: GSS: Em:AQ279830"
misc_feature 16104..16193
/note="match: GSS: Em:AQ074332"
complement(16881..16954)
/gene="dJ581P3.1"
/note="match: GSS: Em:AQ975997"
/pseudo
misc_feature 16980..17146
/note="match: GSS: Em:AQ373489"
misc_feature 17150..17265
/note="match: GSS: Em:AQ373489"
repeat_region 17389..17749
/note="L2 repeat: matches 1140..1536 of consensus"
repeat_region 17859..18030
/note="L2 repeat: matches 1791..1950 of consensus"
repeat_region 18405..18442
/note="19 copies 2 mer gt 97% conserved"
complement(19050..19224)
/note="match: GSS: Em:AQ472547"
repeat_region 19051..19092
/note="21 copies 2 mer ag 90% conserved"
complement(19093..19250)
/note="match: GSS: Em:AQ271994"
repeat_region 19142..19175

```

Alignment Scores:  
Pred. No.: 1.62e-30 Length: 73666

```

Score: 44.00 Matches: 44
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.08% Indels: 0
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x HSDJ581P3 (1-73666)
QY 202 PheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValValAla 221
Db 71995 TTCAGTGGCCCTCATTTCTCTGAGAGATGGCAATGAGACTGCTCCCTGAGGTGTGCC 72054
QY 222 ThrSerGlyTyrAlaLeuLeuHisPheSerAspAlaAlaTyrAsnLeuThrGlyPhe 241
Db 72055 ACATCAGGTATGCTTGTCTGTCGTCATTTTATAGTATGCTCTTATATTTGACTGGATT 72114
QY 242 AsnIleThrTyr 245
Db 72115 AATATTACTTAC 72126

RESULT 68
AF273610 237 bp mRNA linear MAM 09-JUL-2000
LOCUS Sus scrofa mahogany mRNA, partial cds.
DEFINITION AF273610
ACCESSION AF273610
VERSION AF273610.1 GI:89980838
KEYWORDS Sus scrofa (pig)
SOURCE Sus scrofa
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 237)
AUTHORS Matteri,R.L. and Dyer,C.J.
TITLE Sus scrofa mahogany mRNA
JOURNAL Unpublished
TITLE 2 (bases 1 to 237)
AUTHORS Matteri,R.L. and Dyer,C.J.
JOURNAL Direct Submission
TITLE Submitted (01-JUN-2000) Animal Physiology Research Unit, USDA -
JOURNAL Agricultural Research Service, University of Missouri, 920 East
Campus Dr., Room S-143 ASRC, Columbia, MO 65211, USA

FEATURES
source
1..237
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="brain: hypothalamus"
<1..>237
/function="possible role in appetite control and energy
regulation"
/note="transmembrane protein"
/codon_start=2
/product="mahogany"
/protein_id="AAF82297.1"
/db_xref="GI:8980839"
/translation="TDCVSTANTNDCHNCNDHCPNMHSCSTGQISIAKYENCPKDN
PMYCNKKTSCRSCLDQNCQWPRNQECIALPE"

CDS
Alignment Scores:
Pred. No.: 1.37e-30 Length: 237
Score: 42.00 Matches: 42
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.94% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AF273610 (1-237)
QY 744 TyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysThrSerCysArg 763
Db 110 TATGAGAAATCCCAAGGATACCCCATGTACTTACTGTACACAGACACGCTGCGG 169

```

```

QY 764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783
Db 170 AGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCGCCGGAATCAGAGTGCAATGGCCCTG 229

QY 784 ProGlu 785
Db 230 CCCGAA 235

RESULT 69
AB06968S27
LOCUS AB06968S27 5036 bp DNA linear ROD 04-APR-2002
DEFINITION Rattus norvegicus Atrn gene for membrane attractin, exon 29 and
complete cds.
ACCESSION AB049248
VERSION 21143347
KEYWORDS GI:12275389
SEGMENT
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE
1 Kuramoto,T., Kitada,K., Inui,T., Sasaki,Y., Ito,K., Hase,T.,
Kawaguchi,S., Ogawa,Y., Nakao,K., Barsh,G.S., Nagao,M., Ushijima,T.
and Serikawa,T.
2 (bases 1 to 5036)
Kuramoto,T. and Ushijima,T.
Direct Submission
Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center
Research Institute, Carcinogenesis Division, Tsukiji 5-1-1,
Chuo-ku, Tokyo 104-0045, Japan (E-mail:tkuramoto@ncc.ncc.go.jp,
Tel:81-33542-2511, Fax:81-5565-1753)
FEATURES
Location/Qualifiers
1..5036
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/chromosome="3"
join(AB069686.1:3755..4173,AB069687.1:2348..2431,
AB049224.1:515..628,AB049225.1:483..611,
AB049226.1:171..1767,AB049227.1:573..663,
AB049228.1:627..870,AB049229.1:519..702,
AB049230.1:496..650,AB049231.1:504..588,
AB049232.1:336..556,AB049233.1:321..442,
AB049234.1:348..486,AB049235.1:795..985,
AB049236.1:488..707,AB049237.1:595..1516,
AB049238.1:445..582,AB049239.1:107..200,
AB049240.1:422..474,AB049241.1:117..211,
AB049242.1:507..585,AB049243.1:366..523,
AB049244.1:508..569,AB049245.1:522..600,
AB049246.1:500..607,AB049247.1:548..662,544..5036)
/gene="Atrn"

CDS
join(AB069686.1:3755..4173,AB069687.1:2348..2431,
AB049224.1:515..628,AB049225.1:483..611,
AB049226.1:171..1767,AB049227.1:573..663,
AB049228.1:627..870,AB049229.1:519..702,
AB049230.1:496..650,AB049231.1:504..588,
AB049232.1:336..556,AB049233.1:321..442,
AB049234.1:348..486,AB049235.1:795..985,
AB049236.1:488..707,AB049237.1:595..1516,
AB049238.1:445..582,AB049239.1:107..200,
AB049240.1:422..474,AB049241.1:117..211,
AB049242.1:507..585,AB049243.1:366..523,
AB049244.1:508..569,AB049245.1:522..600,
AB049246.1:500..607,AB049247.1:548..662,544..5036)
/gene="Atrn"

```

```

/codon_start=1
/product="membrane attractin"
/protein_id="BAB21058.1"
/db_xref="GI:12275390"
/translation="MVAAAABAETAEARLGRVYTATAAPAGMKERHQRCAATCAWRP
WPRAGLCPLRVLSRALSPPELPLPLLSLILLPLPREAEEAAYAAVAVSGSAAAEAK
ECURPCVNGRCNPGTCVCYPTGWVGEQCGHGRFLRTGSSGFVTDGPNRYKTK
CTWLEQPNKIMRLRFNHPATECSWDHLYVPGDSIYAPLIAAFSLIYPERDNET
APEVTVTSYALLHFFSDAAYNLTFNITFDMCPNPNCGRGECSNSSSVCECSC
SENWGSCECDIPHCTDNCDFPHRGICNASDTGSCGFPHWQPGCSIPVPANOSFWTR
REYSDLKIPRASHKAVVNGIMVVGGMFNHSDYSMLAYDLASRWLSLINSVSV
VVRGSHLGHDKIYMGKIDSTNVNLEVRHILNESWLLTPKADQYAVVGH
SAHIVTSSGRVNLVIFGCPHYGISVQVEDLXNTWSILQTGALVQGGYGHSS
VYDHRTKALYVHGYSKAFKANKRLADLLRYHVDITQMTLLKDSRFRFLYHIAVIS
GTWLVFGCNTNDTSMHGAFCFSDPMAYDIACDRWSVLPRLHLDVNRFGHSAVL
HNATMYVFGFNLSILLSDVLVFTSEQDHAHSEAAACVAGFIRICLMDTQSRCTSE
LHSTBEQAEKLKSECFKRTLDHRCDOHTDCYSTANTNDCHWNCNCHVPNHSCTSG
QISTAKYDNCPKDNMYCNKTKSCSALDQNCOWEPNQEICIALPENTCIGLHWLV
GNSCLIKTAKENVDAKLSRNHNAFLASLTQKKYFVFLKOLELWQSQSKLTL
TPWVGLKXINVTWCHEDMFPFNTSLLOWMPSPSDAGFCGLISESTRGKAKATCIN
PLGSCVERPANHSKQCRTPCALRTACGECTSSSECMWCSNNKQCVDSNAYVAFV
FGQMEWYTWSSCPPENCSCYCTCSHLEQFGCGWCTPDSNTGKGCIEGYSKGFVPM
PSHASTGNVYQPLLNSMCLSDSRYNWSPHCPACQNGHSGKINGISICEKEDLTT
GKHCTCISGFGPTNGKQCPCKNGHASLGNNTNGKCTCTTGKVGECQCEVE
NRYQGNLKTCTCYVTLIDYQFTFSLQSDRDYVTAINFVATPDEONRDLDMFINASK
NFENLITWASFPAGTGTGEVVPVSKTNLKEYKDSFNSEKEDFRNHPNITFEVYVSN
FTPIKIQIAFSQHSNFMDLVQFVTFPFCFLSILLVAAVWIKQSCWASRRRQLL
REMQQMASRFPASVNVVALETDEEPPDLIGSITKVPKPIALEPCFGKNAALVSVFVFL
PRGLGGTTPPGQSLAVASALVDISQQMPIVYKEKSGAVNRKQPPAQPGTCTI"
544..5036
/gene="Atrn"
/product="membrane attractin"
/number=29

ORIGIN
Alignment Scores:
Pred. No.: 1,46e-28 Length: 5036
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.87% Indels: 0
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x AB06968S27 (1-5036)

QY 1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrIys 1408
Db 543 GGTCTTGCTGGCCAGTGCCTCGTGACATTTCTCAGCAGATGCCAATAGTGTAAG 602

QY 1409 GlulysSerGlyAlaValArgAsnArgIysGlnGlnProProAlaGlnProGlyThrCys 1428
Db 603 GAGAAGTCAGGAGCTGTAAAGAACCCGGAAGCAGCAGCCCCCTGCACAGCTGGAACCTGC 662

QY 1429 Ile 1429
Db 663 ATT 665

RESULT 70
H009384S27
LOCUS H009384S27 5790 bp DNA linear PRI 14-DEC-2000
DEFINITION Homo sapiens attractin precursor (ATRn) gene, exon 30 and complete
cde, alternatively spliced.
ACCESSION AF218915
VERSION AF218915.1 GI:8118081
KEYWORDS
SEGMENT 27 of 27
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE
1 (bases 1 to 5790)
Tang,W., Gunn,T.M., McLaughlin,D.F., Barsh,G.S., Schloeman,S.F.
and Duke-Cohan,J.S.

```



**TITLE** Secreted and membrane attractin result from alternative splicing of the human ATRN gene  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)  
**MEDLINE** 20283944  
**PUBMED** 10811918  
**REFERENCE** 2 (bases 1 to 5790)  
**AUTHORS** Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S., Schlossman, S.F. and Duke-Cohan, J.S.  
**TITLE** Direct Submission  
**JOURNAL** Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS, Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115, USA

**FEATURES**  
 source

Location/Qualifiers  
 1. .5790  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20p13"  
 order(AF218889.1:745..1576,AF218890.1:1..3711,  
 AF218891.1:1..3058,AF218892.1:1..1152,AF218893.1:1..1232,  
 AF218894.1:1..1032,AF218895.1:1..1575,AF218896.1:1..2018,  
 AF218897.1:1..3038,AF218898.1:1..1552,AF218899.1:1..3589,  
 AF218900.1:1..1267,AF218901.1:1..3112,AF218902.1:1..1728,  
 AF218903.1:1..3257,AF218904.1:1..1674,AF218905.1:1..1675,  
 AF218906.2:1..966,AF218907.1:1..1781,AF218908.1:1..1704,  
 AF218909.1:1..1976,AF218910.1:1..1364,AF218911.1:1..2214,  
 AF218912.1:1..1556,AF218913.1:1..1592,AF218914.1:1..1394,  
 1..5060)  
 /gene="ATRN"

gene

join(AF218889.1:745..1233,AF218890.1:1:2877..2960,  
 AF218891.1:810..923,AF218892.1:818..946,  
 AF218893.1:621..826,AF218894.1:529..697,  
 AF218895.1:1343..1433,AF218896.1:926..1169,  
 AF218897.1:1114..1297,AF218897.1:2047..2201,  
 AF218898.1:659..743,AF218899.1:1585..1805,  
 AF218900.1:728..849,AF218901.1:440..578,  
 AF218902.1:2163..2353,AF218902.1:533..752,  
 AF218903.1:1441..1626,AF218903.1:2189..926,  
 AF218904.1:810..947,AF218905.1:833..926,  
 AF218906.2:49..101,AF218907.1:802..896,  
 AF218908.1:873..951,AF218909.1:392..549,  
 AF218911.1:1348..1409,AF218912.1:748..826,  
 AF218913.1:747..854,AF218914.1:735..849,660..5060)  
 /gene="ATRN"

mRNA

join(AF218889.1:745..1233,AF218890.1:1:2877..2960,  
 AF218891.1:810..923,AF218892.1:818..946,  
 AF218893.1:621..826,AF218894.1:529..697,  
 AF218895.1:1343..1433,AF218896.1:926..1169,  
 AF218897.1:1114..1297,AF218897.1:2047..2201,  
 AF218898.1:659..743,AF218899.1:1585..1805,  
 AF218900.1:728..849,AF218901.1:440..578,  
 AF218902.1:2163..2353,AF218902.1:533..752,  
 AF218903.1:1441..1626,AF218903.1:2189..926,  
 AF218904.1:810..947,AF218905.1:833..926,  
 AF218906.2:49..101,AF218907.1:802..896,  
 AF218908.1:873..951,AF218909.1:392..549,  
 AF218911.1:1348..1409,AF218912.1:748..826,  
 AF218913.1:747..854,AF218914.1:735..849,660..5060)  
 /gene="ATRN"

CDS

/product="membrane attractin precursor"  
 join(AF218889.1:824..1233,AF218890.1:2877..2960,  
 AF218891.1:810..923,AF218892.1:818..946,  
 AF218893.1:621..826,AF218894.1:529..697,  
 AF218895.1:1343..1433,AF218896.1:926..1169,  
 AF218897.1:1114..1297,AF218897.1:2047..2201,  
 AF218898.1:659..743,AF218899.1:1585..1805,  
 AF218900.1:728..849,AF218901.1:440..578,  
 AF218902.1:2163..2353,AF218902.1:533..752,  
 AF218903.1:1441..1626,AF218903.1:2189..926,  
 AF218904.1:810..947,AF218905.1:833..926,  
 AF218906.2:49..101,AF218907.1:802..896,  
 AF218908.1:873..951,AF218909.1:392..549,  
 AF218911.1:1348..1409,AF218912.1:748..826,  
 AF218913.1:747..854,AF218914.1:735..849,660..784)  
 /gene="ATRN"  
 /note="Method: conceptual translation with partial peptide sequencing"  
 /codon\_start=1  
 /product="membrane attractin precursor"  
 /protein\_id="AA72881.1"  
 /db\_xref="GI:8118082"

SNLKLPRASHKAVVNGNMVVGGMFPHNSDYNMVLAYDLASREWLPLNRVNNVVVR  
 YGSLALYKDKIMYGCKIDYTGNTVNELEFPHHNSWVLLTPFKAQYGVYGHSAH  
 IVTLKNGRVVLMVIFGHCPIDYGVINVOEYDLDDKNTWSILHTQCALVQGGYGVHSSVVD  
 HRTALVHGQYKAFSANKYRLADLVYVDVDTQMTILKDSREFFLYLPAVTVSGTM  
 LVFGNTHNTSMHGKACFFSDPMAYDACADWSVLPRPDLHDDHNRFGHSAVLHNS  
 TMVYGGFNSLLSDILVTFSEQCDARHSEACAAAGPGIRCVWNTGSSQCSISWALAT  
 DBEELKSECFKRLTDHDCDOHTDCYANTINDCHNDHCVPRHNSHSEGGQIS  
 IFREYCPKDNPMYICNKTSCRSALDQNCQWEPNQSCIALPENICGHWLVGNS  
 CLKTITAKENYDKLPCRNHNAALASLTQKKVEFLKQLRIMQSQSKLITLTPW  
 VLGRTKINSVYCWEDMSPFTNSLLOWPPEPSDAGFCILSEPSSTGLKAATCINPLN  
 GSVCEPANHSAKQCTPCALRTACGDCGSGSECCWCSNMKOCVDSNAVASFPFGQ  
 CMWYVTPSTCPPEKSCYCTCHLQPGCGWCTDPSNTGKGCISGKYGKVPMPGQ  
 APTGNFYQPLLNSMCLEDSRYNWSFIFHCPCACQNGHSHKCNQSCICEKENLITGKH  
 CETCISGYGDPNTGGKQCPCKNGHSLCNTNTGKCTTGTGKVGDEQOLCEVKNRY  
 QGNPLRTCYTLLIDYQTFSLQKDDRYTAINFVATPEQNRDLDMFINASKNFY  
 LNIITWASFSAGTQAGEEMPWSTKNIKEYKDSFSLVAAVWKIKOSWASRRREQLREM  
 PIKIQIAFSOHSNFMDLVQFFVTFSCFLSLVAAVWKIKOSWASRRREQLREM  
 QMASRPASVNVALSTDEPDLIGGSIKTPKPIALEPCFNGKAAVLSVFLVRLPRG  
 LGIIPPGQSGLAVALVDISQMPVYKEKSGAVNRKQPPAQPGTICI"  
 join(AF218889.1:1073..1233,AF218890.1:2877..2960,  
 AF218891.1:810..923,AF218892.1:818..946,  
 AF218893.1:621..826,AF218894.1:529..697,  
 AF218895.1:1343..1433,AF218896.1:926..1169,  
 AF218897.1:1114..1297,AF218897.1:2047..2201,  
 AF218898.1:659..743,AF218899.1:1585..1805,  
 AF218900.1:728..849,AF218901.1:440..578,  
 AF218902.1:2163..2353,AF218902.1:533..752,  
 AF218903.1:1441..1626,AF218903.1:2189..926,  
 AF218904.1:810..947,AF218905.1:833..926,  
 AF218906.2:49..101,AF218907.1:802..896,  
 AF218908.1:873..951,AF218909.1:392..549,  
 AF218911.1:1348..1409,AF218912.1:748..826,  
 AF218913.1:747..854,AF218914.1:735..849,660..781)  
 /gene="ATRN"  
 /product="membrane attractin"  
 660..5060  
 /gene="ATRN"  
 /number=30  
 5038..5043  
 /gene="ATRN"

ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,63e-28 Length: 5790  
 Score: 41.00 Matches: 41  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.87% Indels: 0  
 DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x H009384S27 (1-5790)

QY 1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLys 1408  
 Db 659 GGTCTTCTGTGGCAGCGCCCTGGTGGACATTTCTCAGCAGAGTCCGATGTACAG 718  
 QY 1409 GlyLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCys 1428  
 Db 719 GAGAAGTCAGGAGCGGTGAGAAACCGGAGCAGCAGCCCCCTGCACAGCCTGGGACCTGC 778  
 QY 1429 Ile 1429  
 Db 779 ATC 781

RESULT 71  
 MMWANNY2  
 LOCUS MMWANNY2 7437 bp DNA linear ROD 31-MAR-1999  
 DEFINITION Mus musculus mahogany (mg) gene, exons and partial cds.  
 ACCESSION AF120318  
 VERSION AF120318.1 GI:4545288  
 KEYWORDS 2 of 2  
 SEGMENT Mus musculus (house mouse)  
 SOURCE

```

ORGANISM      Mus musculus
REFERENCE     Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS       Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
TITLE         1 (bases 1 to 7437)
JOURNAL       Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
MEDLINE       Schlossman,S.F., Duke-Cohan,J.S. and Bash,G.S.
PUBMED       99184160
FEATURES      The mouse mahogany locus encodes a transmembrane form of human
SOURCE        attractin
JOURNAL       Nature 398 (6723), 152-156 (1999)
MEDLINE       99184160
PUBMED       10086356
REFERENCE     2 (bases 1 to 7437)
AUTHORS       Gunn,T.M., Miller,K.A., He,L., Hyman,R.W., Davis,R.W., Azarani,A.,
TITLE         Direct Submission
JOURNAL       Submitted (14-JAN-1999) HHMI, Stanford, Beckman Center B271,
JOURNAL       Stanford, CA 94305-5323, USA
FEATURES      Location/Qualifiers
SOURCE        1..7437
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10090"
              /chromosomes="2"
              /clone="BAC389B9"
              order(AFI20317.1:1.<1..9085,1..>7437)
              /gene="mg"
              join(AFI20317.1:1.<1870..1931,AFI20317.1:6879..6957,
              282..389,2342..2456,5156..>7437)
              /gene="mg"
              /product="mahogany"
              join(AFI20317.1:1.<1870..1931,AFI20317.1:6879..6957,
              282..389,2342..2456,5156..5280)
              /gene="mg"
              /note="corresponding cDNA sequence found in GenBank
              Accession Number AF119821"
              /codon_start=1
              /product="mahogany"
              /protein_id="AAD22476.1"
              /db_xref="GI:4545290"
              /translation="TAFSOSHFMDLVOFFVFSCFLSLLLVAAVVWKIQSCWASR
              REQLLRQMOASPPASVNVVALETDEPPDLGSIKTKPKPALEPCGNKAAYL
              SVFVRLPGLGIPFGPGSGLAVASALVDISQMFIVYKESGAVRNKQPPAPQGT
              C."
              exon       282..389
              /gene="mg"
              exon       2342..2456
              /gene="mg"
              exon       5256..57437
              /gene="mg"
              3'UTR     5281..>7437
              /gene="mg"
ORIGIN
Alignment Scores:
Pred. No.:    1.99e-28      Length:      7437
Score:        41.00        Matches:    41
Percent Similarity: 100.00%  Conservative: 0
Best Local Similarity: 100.00%  Mismatches: 0
Query Match:   2.87%        Indels:    0
DB:           10          Gaps:      0
US-09-787-097-12 (1-1429) x WMAHNY2 (1-7437)
QY 1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLys 1408
Db 5155 GGTCTCGCTGTGCCAGTGCCTCGTGTGCACATTCTCAGCAGATGCACATAGTGATACAAAG 5214
QY 1409 GluIysSerGlyAlaValAlaArgenArgLysGlnGlnProProAlaGlnProGlyThrCys 1428
Db 5215 GAGAGTCTAGGAGCTGTAGAAACCGAGACAGACGCCCTGCACAGCTCGAACCCTGC 5274
QY 1429 Ile 1429
|||

```

```

Db 5275 ATT 5277
RESULT 72
AC017113/c
LOCUS    AC017113      206329 bp      DNA      linear      HTG 07-JUL-2000
DEFINITION Homo sapiens chromosome 20 clone RP11-574H7, WORKING DRAFT
SEQUENCE, 31 unordered pieces.
ACCESSION AC017113.4 GI:8570339
VERSION    HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS   Homo sapiens (human)
SOURCE     Homo sapiens
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 206329)
           Waterston,R.H.
           The sequence of Homo sapiens clone
           Unpublished
           2 (bases 1 to 206329)
           Waterston,R.H.
           Direct Submission
           Submitted (09-DEC-1999) Genome Sequencing Center, Washington
           University School of Medicine, 4444 Forest Park Parkway, St. Louis,
           MO 63108, USA
COMMENT    On Jun 17, 2000 this sequence version replaced gi:6855220.
           ----- Genome Center -----
           Center: Washington University Genome Sequencing Center
           Center code: WUGSC
           Web site: http://genome.wustl.edu/gsc/index.shtml
           ----- Project Information -----
           Center project name: H NH0574H07
           ----- Summary Statistics -----
           Sequencing vector: M13; 91%
           Chemistry: Dye-terminator; 91% of reads
           Chemistry: Dye-terminator Big Dye; 9% of reads
           Assembly program: Phrap; version 0.990319
           Consensus quality: 184921 bases at least Q40
           Consensus quality: 192397 bases at least Q30
           Consensus quality: 195862 bases at least Q20
           Insert size: 237000; agarose-fp
           Insert size: 203323; sum-of-contigs
           Quality coverage: 3.84 in Q20 bases; agarose-fp
           Quality coverage: 3.66 in Q20 bases; sum-of-contigs
           -----
           * NOTE: This is a 'working draft' sequence. It currently
           * consists of 31 contigs. The true order of the pieces
           * is not known and their order in this sequence record is
           * arbitrary. Gaps between the contigs are represented as
           * runs of N, but the exact sizes of the gaps are unknown.
           * This record will be updated with the finished sequence
           * as soon as it is available and the accession number will
           * be preserved.
           *
           * 1539: contig of 1539 bp in length
           * 1639: gap of unknown length
           * 1640
           * 3259: contig of 1620 bp in length
           * 3260
           * 3359: gap of unknown length
           * 3360
           * 5400: contig of 2041 bp in length
           * 5401
           * 5500: gap of unknown length
           * 5501
           * 7783: contig of 2283 bp in length
           * 7784
           * 7883: gap of unknown length
           * 7884
           * 11241: contig of 3358 bp in length
           * 11242
           * 11341: gap of unknown length
           * 11342
           * 13482: contig of 2141 bp in length
           * 13483
           * 13582: gap of unknown length
           * 13583
           * 16915: contig of 3333 bp in length
           * 16916
           * 17015: gap of unknown length
           * 17016
           * 21346: contig of 4331 bp in length
           * 21347
           * 21446: gap of unknown length
           * 21447
           * 25103: contig of 3657 bp in length
           * 25104
           * 25203: gap of unknown length
           * 25204
           * 28415: contig of 3212 bp in length
           *

```

```

* 28416 28515: gap of unknown length
* 32873: contig of 4358 bp in length
* 32973: gap of unknown length
* 32974 36496: contig of 3523 bp in length
* 36497 36596: gap of unknown length
* 36597 40649: contig of 4053 bp in length
* 40650 40750: gap of unknown length
* 40750 44878: contig of 4129 bp in length
* 44878: gap of unknown length
* 44979 50453: contig of 5375 bp in length
* 50453: gap of unknown length
* 50454 55946: contig of 5493 bp in length
* 55947 56046: gap of unknown length
* 56047 61420: contig of 5374 bp in length
* 61421 61520: gap of unknown length
* 61521 67819: contig of 8299 bp in length
* 67820 72327: contig of 4408 bp in length
* 72328 72427: gap of unknown length
* 72428 77293: contig of 4866 bp in length
* 77294 82636: contig of 5233 bp in length
* 82637 82726: gap of unknown length
* 82727 90186: contig of 7460 bp in length
* 90187 90286: gap of unknown length
* 90287 96584: contig of 6298 bp in length
* 96585 96684: gap of unknown length
* 96685 106019: contig of 9335 bp in length
* 106020 106119: gap of unknown length
* 106120 113530: contig of 7411 bp in length
* 113531 113630: gap of unknown length
* 113631 125151: contig of 11521 bp in length
* 125152 125251: gap of unknown length
* 125252 137677: contig of 12426 bp in length
* 137678 137777: gap of unknown length
* 137778 151263: contig of 13486 bp in length
* 151264 151363: gap of unknown length
* 151364 165370: contig of 14007 bp in length
* 165371 165470: gap of unknown length
* 165471 184704: contig of 19234 bp in length
* 184705 184804: gap of unknown length
* 184805 206329: contig of 21525 bp in length.

```

## FEATURES

```

source
1. 206329
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="20"
/clone="RP11-574H7"

```

## ORIGIN

```

Alignment Scores:
Pred. No.: 2,86e-27 Length: 206329
Score: 41.00 Matches: 41
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.87% Indels: 0
DB: 2 Gaps: 0

```

```

US-09-787-097-12 (1-1429) x AC017113 (1-206329)
Qy 1389 GlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetProIleValTrLys 1408
Db 159327 GGTCTTCTGTGGCCAGCCCTGGTGGACATTTCTCAGCAGATGCGGATAGTACAAAG 159268
Qy 1409 GluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCys 1428
Db 159267 GAGAAGTCAGAGCCGTGAGAAACCGAGACAGCCCCCTTCACAGCCTGGACCTGC 159208
Qy 1429 Ile 1429
Db 159207 ATC 159205

```

## RESULT 73

```

AF415225/25/c
LOCUS
DEFINITION
Mus musculus chromosome 2 clone CITB BAC 38989 strain 129, ***
SEQUENCING IN PROGRESS ***, 25 unordered pieces.
ACCESSION
AF415225
VERSION
AF415225.1 GI:15741219
KEYWORDS
HTG; HTGS PHASE1
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 223026)
Gunn,T.M., Azarani,A., Kim,P.H., Hyman,R.W., Davis,R.W. and
Barsh,G.S.
Identification and preliminary characterization of mouse Adam33
BMC Genet. 3 (1), 2 (2002)
2 (bases 1 to 223026)
Gunn,T.M., Azarani,A., Kim,P., Hyman,R.W., Davis,R.W. and
Barsh,G.S.
Direct Submission
Submitted (30-AUG-2001) HMI, Stanford University, Beckman Center,
Stanford University Medical Center, Stanford, CA 94305, USA
Sequenced to 6x coverage. Clone contains genes: Atrn, Adam33, Sn,
Cdc25b, Cenph, Gfra4 and EST clusters represented by IMAGE clones
3868736, 388220, 775311, 614556, 427645, 807697 and 514017.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
1 764: contig of 764 bp in length
* 765 864: gap of unknown length
* 765 865: contig of 867 bp in length
* 1732 1831: gap of unknown length
* 1832 3030: contig of 1199 bp in length
* 3031 3130: gap of unknown length
* 3131 4476: contig of 1346 bp in length
* 4477 4576: gap of unknown length
* 4577 5597: contig of 1021 bp in length
* 5598 5697: gap of unknown length
* 5698 7086: contig of 1389 bp in length
* 7087 7187: gap of unknown length
* 7188 8614: contig of 1428 bp in length
* 8615 8714: gap of unknown length
* 8715 10700: contig of 1986 bp in length
* 10701 10800: gap of unknown length
* 10801 13233: contig of 2433 bp in length
* 13234 13333: gap of unknown length
* 13334 16153: contig of 2820 bp in length
* 16154 16253: gap of unknown length
* 16254 20801: contig of 4548 bp in length
* 20802 20901: gap of unknown length
* 20902 23601: contig of 2700 bp in length
* 23602 23701: gap of unknown length
* 23702 27816: contig of 4115 bp in length
* 27817 27916: gap of unknown length
* 27917 37125: contig of 9209 bp in length
* 37126 37226: gap of unknown length
* 37226 44951: contig of 7726 bp in length
* 44952 45051: gap of unknown length
* 45052 54136: contig of 9085 bp in length
* 54137 54237: gap of unknown length
* 54237 61674: contig of 7437 bp in length
* 61674 61773: gap of unknown length
* 61773 71077: contig of 9304 bp in length
* 71077 71177: gap of unknown length
* 71177 81674: contig of 10497 bp in length
* 81674 81774: gap of unknown length

```

\* 81775 92951: contig of 11177 bp in length  
 \* 92952 93051: gap of unknown length  
 \* 93052 107739: contig of 14688 bp in length  
 \* 107740 107840: gap of unknown length  
 \* 107840 126356: contig of 18417 bp in length  
 \* 126356 126357: gap of unknown length  
 \* 126357 145400: contig of 19044 bp in length  
 \* 145401 145500: gap of unknown length  
 \* 145501 175039: contig of 29539 bp in length  
 \* 175040 175139: gap of unknown length  
 \* 175140 223026: contig of 47887 bp in length.

## FEATURES

source

1. .223026  
 /organism="Mus musculus"  
 /mol\_type="genomic DNA"  
 /strain="129"  
 /db\_xref="taxon:10090"  
 /chromosome="2"  
 /clone="C1B BAC 389B9"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,04e-27 Length: 223026  
 Score: 41.00 Matches: 41  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.87% Indels: 0  
 DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x AF415225 (1-223026)

QY 1389 GlyLeuAlaValAlaSerIleuValAspIleSerGlnGlnMetProIleValTyrIys 1408  
 Db 56519 GGTCTCGTGGCCAGTCCCTGGTGGACATTTCTCACACATGCCAATAGTGACAG 56460

QY 1409 GluLysSerGlyAlaValAlaArgAsnArgLysGlnGlnProAlaGlnProGlyThrCys 1428  
 Db 56459 GAGAGTCAGGAGCTGTAGAACCGAGAGCAGCCGCTGCACAGCTGGACCTGC 56400

QY 1429 Ile 1429

Db 56399 ATT 56397

## RESULT 74

H009384S12  
 LOCUS H009384S12 1267 bp DNA linear PRI 14-DEC-2000  
 DEFINITION Homo sapiens attractin precursor (ATRN) gene, exon 13.  
 ACCESSION AF218900  
 VERSION AF218900.1 GI:8118066

## KEYWORDS

12 of 27

## SEGMENT

Homo sapiens (human)

## SOURCE

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

## REFERENCE

1 (bases 1 to 1267)

Tang, W., Gunn, T.M., McLaughlin, D.F., Barsh, G.S., Schlossman, S.F.

and Duke-Cohan, J.S.

Secreted and membrane attractin result from alternative splicing of

the human ATRN gene

Proc. Natl. Acad. Sci. U.S.A. 97 (11), 6025-6030 (2000)

## PUBMED

20283944

10811918

## AUTHORS

2 (bases 1 to 1267)

Tang, W., Gunn, T.M., McLaughlin, D.F., Freeman, G.J., Barsh, G.S.,

Schlossman, S.F., and Duke-Cohan, J.S.

Direct Submission

Submitted (27-DEC-1999) Department of Cancer Immunology and AIDS,

Dana-Farber Cancer Institute, 44 Binney Street, Boston, MA 02115,

USA

## FEATURES

source

1. .1267  
 /organism="Homo sapiens"

/mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="20"  
 /map="20p13"  
 728..849  
 /gene="ATRN"  
 /number=13

exon

## ORIGIN

Alignment Scores:  
 Pred. No.: 4.43e-28 Length: 1267  
 Score: 40.00 Matches: 40  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 2.80% Indels: 0  
 DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x H009384S12 (1-1267)

QY 699 LeuAspHisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsn 718  
 Db 730 CTTGACCATGACATGTGACACACACACATTTGTACAGCTGCACAGCCACACCAAT 789

QY 719 AspCysHisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGln 738  
 Db 790 GACTGCCACTGGTGCATGACCATTTGTCTCCAGGAACACACAGCTGCTCAGAGGCCAG 849

## RESULT 75

AB06968S13

## LOCUS

AB06968S13

1071 bp DNA linear ROD 04-APR-2002

DEFINITION Rattus norvegicus Atrn gene for membrane attractin, exon 14.

ACCESSION AB049234

VERSION AB049234.1 GI:12275375

KEYWORDS 13 of 27

SEGMENT

Rattus norvegicus (Norway rat)

## ORGANISM

Rattus norvegicus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Muridae; Murinae;

Rattus.

## REFERENCE

1

Kawaguchi, S., Ogawa, Y., Nakao, K., Barsh, G.S., Nagao, M., Ushijima, T.

and Serikawa, T.

Attractin/mahogany/zitter plays a critical role in myelination of

the central nervous system

Proc. Natl. Acad. Sci. U.S.A. 98 (2), 559-564 (2001)

## PUBMED

21143347

11209055

## REFERENCE

2 (bases 1 to 1071)

Kuramoto, T. and Ushijima, T.

Direct Submission

Submitted (23-SEP-2000) Takashi Kuramoto, National Cancer Center

Research Institute, Carcinogenesis Division; Tsukiji 5-1-1,

Chuo-Ku, Tokyo 104-0045, Japan [E-mail:tkuramoto@gan2.ncc.go.jp,

Tel:81-33542-2511, Fax:81-5565-1753]

Location/Qualifiers

1. .1071

/organism="Rattus norvegicus"

/mol\_type="genomic DNA"

/db\_xref="taxon:10116"

/chromosome="3"

348..486

/gene="Atrn"

/product="membrane attractin"

/note="CDS is reported in Acc# AB049248"

/number=14

## ORIGIN

Alignment Scores:

Pred. No.: 3.56e-27 Length: 1071

Score: 39.00 Matches: 39

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.73% Indels: 0  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x AB06968813 (1-1071)

Qy	746	AsnCysProLysAspAsnProMetTyrCysAsnLysLysThrSerCysArgSerCys	765
Db	369	AATTGCCCCCAAGGATTAACCCCATGTACTACTGCATAAGAAAAACCACTGCAGGAGCTGT	428
Qy	766	AlaLeuAspGlnAsnCysGlnTTPGluProArgAsnGlnGluCysTleAlaLeuPro	784
Db	429	GCCTAGATCAGAAATTGCCAGTGGGAACCTCGAAATCAAGAGTGCATCGCCCTGCC	485

Search completed: March 1, 2004, 12:04:58  
Job time : 15574 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 1, 2004, 02:00:44 ; Search time 1075 Seconds  
(without alignments)  
5647.140 Million cell updates/sec

Title: US-09-787-097-12

Perfect score: 1429

Sequence: 1 MVAAATAAEARLRRTAATA.....KSGAVNRKQPPAQPGTCT 1429

Scoring table: OLIGO

Xgapop 60.0 , Xgapext 60.0

Ygapop 60.0 , Ygapext 60.0

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Word size: 1

Total number of hits satisfying chosen parameters: 6745146

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp  
-Q/cgn2.1/USPTO.spool.p/US09787097/runat 26022004 063014 23737/app\_query.fasta\_1.1607  
-DS=N Geneseq 29Jan04 -CEMT=fastap -SUFFIX=oligo.rng -MINMATCH=0.1 -LOOPCL=0  
-LOPEXT=0 -UNIFS=bits -START=1 -END=1 -MATRIX=oligo -TEAMS=human40.cdi  
-LIST=100 -DOCALIGN=200 -THR\_SCORE=quality -THR\_MIN=1 -ALIGN=75 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09787097 @CEN 1.1.1137@runat 26022004 063014 23737 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : N\_Geneseq\_29Jan04.\*

1: Geneseqn1980s.\*

2: Geneseqn1990s.\*

3: Geneseqn2000s.\*

4: Geneseqn2001as.\*

5: Geneseqn2001bs.\*

6: Geneseqn2002s.\*

7: Geneseqn2003as.\*

8: Geneseqn2003bs.\*

9: Geneseqn2003cs.\*

10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1429	100.0	4290	3	Aaz52276 Human mem
2	1429	100.0	8561	5	Aas72659 DNA encod
3	1267	88.7	3819	3	Aaz52277 Human sol
4	1141	79.8	8590	3	Aaz91920 Human mah
5	1123	78.6	4068	3	Aaz52275 Human mem
6	979	68.5	4072	3	Aaz521921 Human mah
7	874	61.2	3597	3	Aaz52274 Human sol
8	874	61.2	3597	5	Aas72657 DNA encod

9	874	61.2	3597	9	ADD70997	Human att
10	496	34.7	2625	3	Aaz91922	Human mah
11	467	32.7	6373	3	Aaz91919	Human mah
12	289	20.2	3490	5	AAS72660	DNA encod
13	206	14.4	8739	9	ADBS3249	Primary r
14	181	12.7	8827	3	Aaz91916	DNA encod
15	142	9.9	3165	5	AAS78781	Wild type
16	142	9.9	3313	5	AAS76294	DNA encod
17	110	7.7	671	5	AAS72658	DNA encod
18	104	7.3	668	4	AAS02409	Human sec
19	94	6.6	2419	3	Aaz91917	Murine ma
20	92	6.4	481	4	AAL14686	Human bre
21	92	6.4	1012	2	AAX84568	Biorhyth
22	82	5.7	246	6	ABK38928	cDNA enco
23	82	5.7	246	7	ACA11257	Human lun
24	82	5.7	246	7	ACA02443	Lung canc
25	79	5.5	482	4	AAL23548	Human bre
26	77	5.4	207433	5	ABZ72040	Gene 216
27	77	5.4	207433	7	ABX74891	BAC1088L2
28	76	5.3	1051	3	Aaz91918	Murine ma
29	73	5.1	90050	3	Aaz91925	Wild type
30	69	4.8	531	4	AAI43425	Probe #12
31	69	4.8	531	6	ABS11523	Human gen
32	64	4.5	193	4	AAI56405	Probe #25
33	64	4.5	193	6	ABS23881	Human gen
34	64	4.5	292	3	AAA45505	Human sec
35	60	4.2	182	4	AAI24099	Probe #14
36	60	4.2	182	4	ABA69223	Human foe
37	60	4.2	182	4	AAI49392	Probe #18
38	60	4.2	182	4	ABA51213	Human bre
39	60	4.2	182	4	ABA36140	Probe #14
40	60	4.2	182	4	AAK43324	Human bon
41	60	4.2	182	4	AAK17517	Human bra
42	60	4.2	182	4	ABS42949	Human liv
43	60	4.2	182	5	AAI09674	Probe #96
44	60	4.2	182	6	ABS17404	Human gen
45	57	4.0	375	4	AAI14910	Probe #48
46	57	4.0	375	4	ABA56648	Human foe
47	57	4.0	375	4	AAI36262	Probe #49
48	57	4.0	375	4	ABA46100	Human bre
49	57	4.0	375	4	ABA26254	Probe #47
50	57	4.0	375	4	AAK30307	Human bon
51	57	4.0	375	4	AAK04785	Human bra
52	57	4.0	375	4	ABS29950	Human liv
53	57	4.0	375	5	AAI04689	Probe #46
54	57	4.0	375	6	ABS04892	Human gen
55	51	3.6	2267	3	Aaz91926	Wild type
56	49	3.4	2267	5	AAS69062	DNA encod
57	46	3.2	402	4	AAI13789	Probe #37
58	46	3.2	402	4	ABA55499	Human foe
59	46	3.2	402	4	AAI35149	Probe #38
60	46	3.2	402	4	ABA45027	Human bre
61	46	3.2	402	4	ABA25215	Probe #36
62	46	3.2	402	4	AAK29197	Human bon
63	46	3.2	402	4	AAK03733	Human bra
64	46	3.2	402	4	ABS28814	Human liv
65	46	3.2	402	5	AAI03665	Probe #36
66	46	3.2	402	6	ABS03749	Human gen
67	45	3.1	137	4	AAI23009	Probe #12
68	45	3.1	137	4	ABA68099	Human foe
69	45	3.1	137	4	AAI48309	Probe #16
70	45	3.1	137	4	ABA50165	Human bre
71	45	3.1	137	4	ABA35129	Probe #13
72	45	3.1	137	4	AAK42239	Human bon
73	45	3.1	137	4	AAK16486	Human bra
74	45	3.1	137	4	ABS41845	Human liv
75	45	3.1	137	5	AAI08671	Probe #86
76	45	3.1	137	6	ABS16289	Human gen
77	33	2.3	5973	3	Aaz91924	Wild type
78	30	2.1	16956	3	Aaz91923	Wild type
79	29	2.0	4134	7	ABZ58710	Murine ho
80	29	2.0	4140	7	ABZ58709	Human hom
81	29	2.0	6517	4	AAH98748	Human EST

82 22 1.5 657 7 AB258714  
 83 22 1.5 845 6 ABQ52396  
 84 22 1.5 845 6 ABQ52397  
 85 22 1.5 982 6 ABL39784  
 86 22 1.5 1654 7 ABX34730  
 87 22 1.5 3609 6 ABG63434  
 88 22 1.5 5312 6 ABL32197  
 89 20 1.4 60 6 ABN42673  
 90 20 1.4 503 5 AAS77314  
 91 19 1.3 6532 5 AAS86766  
 92 18 1.3 65 6 ABL57242  
 93 17 1.2 51 4 AAL33285  
 94 16 1.1 233 6 ABQ97182  
 95 16 1.1 423 5 AAS90685  
 96 16 1.1 1104 5 AAS90708  
 97 16 1.1 1788 5 AAS70516  
 98 15 1.0 1383 6 AAL54064  
 99 15 1.0 4461 4 ABL14674  
 100 14 1.0 510 5 AAS90687

Abz58714 Upregulat  
 Abq52396 Oligonuc1  
 Abq52397 Oligonuc1  
 Abl39784 Human NS  
 Abx34730 Human mcd  
 Abg63434 Human mcd  
 Abl32197 Human imm  
 Abn42673 Human spl  
 Aas77314 DNA encod  
 Aas86766 DNA encod  
 ABL57242 Mouse spl  
 Aal33285 Human SNP  
 Abq97182 Mouse ES  
 Aas90685 DNA encod  
 Aas90708 DNA encod  
 Aas70516 DNA encod  
 Aal54064 cDNA of h  
 Abl14674 Drosophil  
 Aas90687 DNA encod

## ALIGNMENTS

RESULT 1  
 AAZ52276  
 ID AAZ52276 standard; cDNA; 4290 BP.  
 AC AAZ52276;  
 XX  
 XX 18-JUL-2000 (first entry)  
 XX  
 XX Human membrane attractin-2 cDNA.  
 XX  
 XX Human; membrane attractin-2; immune response; macrophage; monocyte;  
 KW T cell; immunostimulant; immunosuppressed patient; cancer;  
 KW immunodeficiency syndrome; transplant; autoimmune disease; ss.  
 XX  
 XX Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 XX CDS 1..4290  
 XX /\*tag= a  
 XX /product= "Membrane attractin-2"  
 XX  
 XX WO200015651-A1.  
 XX  
 XX 23-MAR-2000.  
 XX  
 XX 14-SBP-1999; 99WO-US020948.  
 XX  
 XX 14-SBP-1998; 98US-0100137P.  
 XX  
 XX (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 XX Duke-Cohan JS, Schlossman SF;  
 XX  
 XX WPI; 2000-271373/23.  
 XX P-PSDB; AAY70691.  
 XX  
 XX Isolated nucleic acids encoding human attractin polypeptides useful for  
 XX enhancing immune responses.  
 XX  
 XX Claim 3; Fig 14; 120pp; English.  
 XX  
 XX The patent discloses four forms of human attractin polypeptides which  
 XX enhance immune response by promoting macrophage and monocyte spreading in  
 XX the presence of T cells. These include soluble attractin-1 and -2 and  
 XX membrane attractin-1 and -2. These various forms of attractin are encoded  
 XX by alternatively spliced mRNA molecule transcribed from a single gene.  
 XX The present sequence is a cDNA encoding human membrane attractin-2.  
 XX Membrane attractin differs from soluble attractin in having a  
 XX transmembrane domain and a cytoplasmic domain at the C-terminal.  
 XX Attractin-2 differs from attractin-1 in having a 74 amino acid insert in

CC the N-terminal. This sequence can be used to enhance immune response in  
 CC immunosuppressed patients such as those undergoing chemo- and radio-  
 CC therapy treatment for cancer or those suffering from common variable  
 CC immunodeficiency syndrome. The proteins may also be used to screen  
 CC modulators (agonists and antagonists) of immune response which may also  
 CC be used to regulate immune reactions. Attractin antibodies can be used to  
 CC inhibit immune response in transplant recipients or patients afflicted  
 CC with autoimmune disease

SQ Sequence 4290 BP; 1060 A; 1025 C; 1120 G; 1085 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0 Length: 4290  
 Score: 1429.00 Matches: 1429  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x AAZ52276 (1-4290)

QY 1 MetValAlaAlaAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAla 20  
 DB 1 ATGTGGCCGCGCGCGCAACTGAGGCAAGGCTGAGGAGGAGCGCGCGCGCGCA 60  
 QY 21 AlaLeuAlaGlyArgSerGlyGlyProHisTirAspTirAspValThrArgAlaGlyArg 40  
 DB 61 GCGCTCGCGGAGGAGCGCGGCGGCGCACTGGGACTGGACCTGACCGGCTGGAGG 120  
 QY 41 ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg 60  
 DB 121 CCGGGGCTGGGGCGCGGCTCGCGCTCCCGCGCTGTGTCTCCACCGCTGCGCGCACGG 180  
 QY 61 LeuLeuLeuLeuLeuLeuLeuProProLeuLeuLeuLeuLeuLeuProCysGlu 80  
 DB 181 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240  
 QY 81 AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu 100  
 DB 241 GCCGAGCCGCGCGCGCGCGCGGCGGTGTCTGGGCTCAGCGCGCGCGCGCGCAAGGAA 300  
 QY 101 CysAspArgProCysValAsnGlyGlyArgCysAsnProClyThrGlyGlnCysValCys 120  
 DB 301 TGTACCGCGCGCTGTGTCAACCGCGGTCGTGTCAACCGCGCGCGCGCGCGCGCGCTGC 360  
 QY 121 ProAlaGlyTirValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly 140  
 DB 361 CCCCGCGCTGGGTGGCGGAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAATCGGA 420  
 QY 141 SerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTir 160  
 DB 421 TCCTCTGGGTGTGTACAGATGGACCTGGAAATATATAATACAAACGAAAGTGCACGTGG 480  
 QY 161 LeuLeuGlyGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu 180  
 DB 481 CTCATTGAGGACAGCCCAATAGAAATATGAGACTTGTTCATCATCTTTGGTACAGAG 540  
 QY 181 CysSerTirAspHisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAla 200  
 DB 541 TGTAGTTGGACCACTTATATATATATATATATGATGGGAGCTCAATTTATGCGCGCTAGT 600  
 QY 201 AlaPheSerGlyLeuLeuValProGluArgAspGlyAsnGluThrValProGluValVal 220  
 DB 601 GCATTTAGTGGCTTCATTGTCTCTGAGAGATGGCAATGAGACTGTCCCTGAGGTGTT 660  
 QY 221 AlaThrSerGlyTyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGly 240  
 DB 661 GCCACATCAGTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720  
 QY 241 PheAsnLeuThrTyrSerPheAspMetCysProCysAsnCysSerGlyArgGlyGluCys 260  
 DB 721 TTTAATATTACTTACAGTTTGTATATGTGTCCAAATAACTGCTCAGCGCGCGAGAGTGT 780



QY 261 LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu 280  
DB 781 AGATCAGTAATAGCAGCGAAACTGTTGAATGTGAATGTTCTGAAACTGGAAAGGTGAA 840  
QY 281 AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsn 300  
DB 841 GCATGTGACATTCCTCACTGTACAGACAACATGTGGTTTTCTTCATCGAGGCATCTGCAAT 900  
QY 301 SerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal 320  
DB 901 TCAAGTGATGTGAGAGATGCTCCCTGCTTCTCAGACTGGCAGGGTCTGGAGTGTTCAGTT 960  
QY 321 ProValProAlaAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuPro 340  
DB 961 CTGTGTACAGCTAACCCAGTCATTTGGACTCGAGAGGAATATCTTAACCTTAAAGCTCCCC 1020  
QY 341 ArgAlaSerHisGlyAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360  
DB 1021 AGAGCATCTCATAAAGCTGTGTCATGGAACATTTATGTGGTTGTGGAGGATATATG 1080  
QY 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380  
DB 1081 TTCACCACTCAGATTATACATGTTCTPAGCGTATGACCTTGCTTCTAGGGAGTGGCTT 1140  
QY 381 ProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyr 400  
DB 1141 CCACTAAACCGTCTCTGCAACAATGTGGTTGTAGATATGGTCATCTTTGGCATTATAC 1200  
QY 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420  
DB 1201 AAGGATAAAATTTACATGTATGGAGGAAAAATGTATCCAACTGGGAATGTGACCAATGAG 1260  
QY 421 LeuArgValPheHisIleHisAsnGlnSerTrpValLeuLeuThrProLysAlaLysGlu 440  
DB 1261 TTGAGAGTTTTTACATTCATTAATGAGTCATGGGTGTTGTTGACCCCTTAAGGCAAGGAG 1320  
QY 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460  
DB 1321 CAGTATGCAAGTGGTGGGCACCTCGCACACATGTTTACACTGAAGAATGGCGAGTGGTC 1380  
QY 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480  
DB 1381 ATGCTGCTCATCTTTGGTCACCTGCCCTCTCTATGATATATGAAGCAATGTGCGAAGATAT 1440  
QY 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGly 500  
DB 1441 GATTTGGATAAGAACACATGAGTATATTACACACCCAGGGTGGCCCTTGTGCAAGGGGT 1500  
QY 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
DB 1501 TAGGGCCATAGCAGTGTTTACGACCATAGGACAGGGCCCTATACGTTTCATGGTGGCTAC 1560  
QY 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAsp 540  
DB 1561 AAGGCTTTTCAGTGCCAAATAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGAT 1620  
QY 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560  
DB 1621 ACCCAGATGTGGACCATCTTCTTCAAGGACAGCCGATTTTCCGTTACTTGCACACACCTGTG 1680  
QY 561 IleValSerGlyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSer 580  
DB 1681 ATAGTCAGTGGAAACCAATGTGTGTTTGGGGAAACACACACATGACACATCTATGAGC 1740  
QY 581 HisGlyAlaLysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600  
DB 1741 CATGGGCCAAATGCTTCTTCTTCAGATTTTCATGGCCCTATGACATTCGCTGTGACCGCTGG 1800  
QY 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620  
DB 1801 TCAGTGTCTCCACAGACCTGATCTCCACCATGATGTCACAGATTTGGCCATTCAGCAGTC 1860  
QY 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIle 640

DB 1861 TTACACACAGCACCACCATGTATGTGTTGGTGGTTTCAATAGTCTCTCTCTCAGGACATC 1920  
QY 641 LeuValPheThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660  
DB 1921 CTGATATTCACTCGGAACAGTGTGATCGCATCGAGTGAAGCCGCTGTGTTAGCAGCA 1980  
QY 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680  
DB 1981 GGACCTGCTAATTCGGTGTGTGTGGAACACAGGGTCTCTCAGTGTATCTGCTGGGCGCTG 2040  
QY 681 AlaThrAspGluGlnGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700  
DB 2041 GCAACTGATGAACAAGAAAAAGTTAAATACAGATGTTTTTCCAAAGAACCTCTGTAC 2100  
QY 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720  
DB 2101 CATGACAGATGTGACCCAGCACAGATTTGTACAGCTGTACAGCCCAACCAATGACTGC 2160  
QY 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
DB 2161 CACTGGTGCATACCACTTGTGTCGCCAGGAACACACAGCTGCTCAGAAGCCAGATCTCC 2220  
QY 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760  
DB 2221 ATTTTATAGGTATGAGATTCGCCCAAGGATAACCCCATGCTACTACTGTAAACAAGAGACC 2280  
QY 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780  
DB 2281 AGCTGACAGAGCTGTGCCCTGGACAGAACTGCCAGTGGAGCCCGGATCAGGAGTGC 2340  
QY 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTyrHisLeuValGlyAsnSerCysLeu 800  
DB 2341 ATTGCCCTTGCCGAAAAATATCTGTGCATTTGGCTGGCATTTGGTTGGAAATCATGTGTTG 2400  
QY 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820  
DB 2401 AAAATTACTACTGCCAAGGAGAAATTATGACAACTGCTAAATGTTCTGTAGAACCAACAT 2460  
QY 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840  
DB 2461 GCCCTTTTGGCTTCTCTTACAAACCCAGAGAAGTAGAATTTGCTCTTAAGCAGCTGCGA 2520  
QY 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860  
DB 2521 ATAAATGACGATCATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGGTGGGCTTCGG 2580  
QY 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880  
DB 2581 AAGATCATGTGTCCTACTGTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTA 2640  
QY 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
DB 2641 CAGTGGATGGCGTCTGAGCCGAGTGTGCTGGATTTCTGTGGAATTTTATCAGAACCCAGT 2700  
QY 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
DB 2701 ACTGGGACCTGAAGGCTGCAACCTGCACTCAACCCATCAATGGTAGTGTCTGTGAAGG 2760  
QY 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
DB 2761 CCTCAAAACCAACAGTCATGAAGCAGTGGCCGACACCATGTGCTTGGAGACAGCATGTGGA 2820  
QY 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960  
DB 2821 GATTGCAACCGCCGACCTCTGAGTGTGATGGTGCAGCAACATGAAGCAGTGTGTGGAC 2880  
QY 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980  
DB 2881 TCCATGCTATGTGGCTCTCTTCCCTTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940  
QY 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnPro 1000

Db 2941 ACCTGCCCTCCGAAATGTTTCAGGCTACTGTACCTGTAGTCAATGCTTGGAGCAACA 3000  
Qy 1001 GlyCysGlyTyrCysThrAspProSerAsnThrGlyLysCysIleGluGlySer 1020  
Db 3001 GGCTGTGGCTGGTGTACTGATCCAGCAATACTGGCAAGGAAATGATAGAGGGTTC 3060  
Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrClyAsnPhetYrProGlnPro 1040  
Db 3061 TATAAAGGACCACTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCAGCC 3120  
Qy 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCys 1060  
Db 3121 CTGCTCAATCCAGCATGTCTAGAGGACAGCAGATACAACCTGGTCTTTTCATCTCACTGT 3180  
Qy 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080  
Db 3181 CCAGCTTGCCATGCAACGCCACACATAATGCAATCAGAGCATCTGTGAGAAGTGT 3240  
Qy 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100  
Db 3241 GAGAACTGACCAAGCAAGCACTGCGAGACTGCAATCTGGCTTCTACGGTATCC 3300  
Qy 1101 ThrAsnGlyCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120  
Db 3301 ACCAATGGAGGAAATGTGAGCCATCAAGTGCATGGGACGCGTCTCTGTGCAACACC 3360  
Qy 1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140  
Db 3361 AACACGGGCAAGTCTCTGACCAACCAAGGGGCTCAGGGGACAGTGCAGTATGT 3420  
Qy 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160  
Db 3421 GAGTAGAAATCGATACCAAGGAAACCTCTCAGAGGAACATGTTATTATCTCTCTTT 3480  
Qy 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrAlaIle 1180  
Db 3481 ATTGACTATCAGTTCACTTACCTTAGCTATATCCAGGAAGATGATCGCTATTACACACTATC 3540  
Qy 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200  
Db 3541 AATTTTGTGGCTACTCTCGAGNACAAACACAGGATTTGACATGTTCAATGCCTCC 3600  
Qy 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220  
Db 3601 AAGAAATTTCAACCTCAACATCACTGGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGA 3660  
Qy 1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240  
Db 3661 GAAGAGATGCTGTTGTTTCAAAAACCAACATTAAGAGATACAAAGATAGTTTCTCTAAT 3720  
Qy 1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260  
Db 3721 GAGAAGTTTGATTTTCGCAACCAACCAATATCACTTCTTCTTCTTCTTCTGCTGCTGCTGCT 3780  
Qy 1261 ThrTrpProIleLysGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280  
Db 3781 ACCTGCCCCATCAAAATTCAGATGGCTTCTCTCAGCAGCAAAATTTTATGACCTGGTA 3840  
Qy 1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValVal 1300  
Db 3841 CAGTCTCTGCTGATTTCTTCAGTCTTCTTCTTCTTCTTCTTCTTCTTCTGCTGCTGCTGCT 3900  
Qy 1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMet 1320  
Db 3901 TGGAAATCAAAAGATGTTGTTGGGCTCCAGACGTAGAGACAATCTTCTCGAGAGATG 3960  
Qy 1321 GlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340  
Db 3961 CAACAGATGGCCAGCGCTCCCTTGGCTCTGTAATGTCGCTTGGAAACAGATGAGGAG 4020  
Qy 1341 ProProAspLeuIleGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360  
Db 4021 CCTCCTGATCTTATTGGGGGGAGTATAAGACTGTTCCCAAAACCCATTCCTGCTGAGCGG 4080

Qy 1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380  
Db 4081 TGTTTTGGCAACAAGCCGCTGCTCTCTGTGTTGTTGTGAGCTCCCTCGAGCCCTGGT 4140  
Qy 1381 GlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer 1400  
Db 4141 GGCATCTCTCTCTCTGGGCACTCAGTCTTCTGTGTGGCAGCCCTGTGGACATTTCT 4200  
Qy 1401 GlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGln 1420  
Db 4201 CAGCAGATCGCATAGTGTACAAGGAGAGTCAAGAGCCGTGAGAAACCGAAGACGACG 4260  
Qy 1421 ProProAlaGlnProGlyThrCysIle 1429  
Db 4261 CCCCTGCACAGCTGGGACCTGCATC 4287  
RESULT 2  
AAS72659  
ID AAS72659 standard; cDNA; 8561 BP.  
XX AAS72659;  
AC AAS72659;  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #8463.  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX Homo sapiens.  
XX WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US0008631.  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
PA Drmanac RT, Liu C, Tang YT;  
PI WPI; 2001-639362/73.  
DR P-PSDB; ABG08472.  
DR  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 8463; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activity. The  
CC polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC responsible for genetic disorders or other traits to assess biodiversity  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS84197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this

CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 8561 BP; 2246 A; 1983 C; 2065 G; 2267 T; 0 U; 0 Other;

```

Alignment Scores:
Pred. No.:      0      Length:      8561
Score:          1429.00 Matches:      1429
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:      100.00% Indels: 0
DB:               5      Gaps: 0

```

US-09-787-097-12 (1-1429) x AAS72659 (1-8561)

Qy	1	MetValAlaAlaAlaThrGluAlaArgLeuArgArgThrAlaAlaThrAla	20
Db	1	ATGGTGGCCGACGGCGGCAACTGAGGCAAGGGCTGAGGAGGAGGAGCGCGGCGACGGCA	60
Qy	21	AlaLeuAglArgSerGlyGlyProHisTrpAspTrpAspValThrArgAlaGlyArg	40
Db	61	GCCTCGCGGGCAGAGCGCGGCGGCACCTGGAGACTGGGACGTGGACACAGGGTGGGAGG	120
Qy	41	ProGlyLeuGlyAlaGlyLeuArgLeuProArgLeuLeuSerProProLeuArgProArg	60
Db	121	CCGGGGCTGGGGCGGGCTGGCGCTCCGGCGGGTGTCTGTCTCCACCGCTGGCGCCACCG	180
Qy	61	LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeuProCysGlu	80
Db	181	CTGCTGCTGCTGTGTGTCTCCGCGCGCGCTGTGTCTGTCTGTCTGTCTGTCTGTGAG	240
Qy	81	AlaGluAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu	100
Db	241	GCCGAGCGCGGGCGGCGCGGGGTGTGGGCTCAGCCGCGAGCGGAGGCGCAAGGAA	300
Qy	101	CysAspArgProCysValAsnGlyArgCysAsnProGlyThrGlyGlnCysValCys	120
Db	301	TGTGACCGGCCCTGTCTCAACGGCGGTGGCTGCACCCCTGGACCGGCAGTGGCTGTC	360
Qy	121	ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGly	140
Db	361	CCGCGCGGTGGGTGGGCGAGCAATGCCAGCACTGCGGGGGCGGCTTCAGACTACTGGA	420
Qy	141	SerSerGlyPheValThrAspGlyProGlyAsnTyrlsTyrlsTyrlsCysThsTrp	160
Db	421	TCCTCTGGGTTGTGACAGTGGACTCGAAATATAAATAACAAAGCAAGTGCACGTGG	480
Qy	161	LeuLeuGluGlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGlu	180
Db	481	CTCATTAAGAGCAGCAATAAGAAATAGACACTTCGTTCAATCAATTTTCTACACAG	540
Qy	181	CysSerTrpAspHisLeuTrValTrAspGlyAspSerIleTyrlsAlaProLeuValAla	200
Db	541	TGTAGTTGGACCATTTATATTATGATGGGACTCAAATTTATGCACCGCTAGTTGCT	600
Qy	201	AlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrValProGluValVal	220
Db	601	GCATTTAGTGGCCCTCAATTTCTCTGAGAGAGATGGCAATGACACTGTCCCTGAGGTTGT	660
Qy	221	AlaThrSerGlyTyrlAlaLeuLeuHisPhePheSerAspAlaAlaTyrlAsnLeuThrGly	240
Db	661	GCCACATCAGTTATGCCTGTCTGCATTTTTTTTAGTGATGTGCTTATATTGACTGGA	720
Qy	241	PheAsnIleThrTyrlSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCys	260
Db	721	TTTAATATATTACTACAGTTTTCATATGTCTCCAAATAACTGCTCAGGCGCGAGGAGAGTGT	780
Qy	261	LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu	280
Db	781	AAGATCATGTAATACGACCGAACTGTGTAATGTGTGAAATGTTCTGAAAACTGAAAAGGTGAA	840
Qy	281	AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisIleArgGlyIleCysAsn	300

Db	841	GCA	GTG	CA	ATT	CCT	CAC	TG	TAC	GAC	CA	ACT	GTG	GGT	TTT	CCT	CA	TG	CG	GC	AT	CG	CA	T	900	
Qy	301	Ser	Ser	Asp	Val	Arg	Gly	Cys	Ser	Cys	Phe	Ser	Asp	Trp	Gln	Gly	Pro	Gly	Cys	Ser	Val				320	
Db	901	TCA	AGT	GAT	GTC	CAG	AG	ATG	CTC	TCT	CTC	AG	ACT	GCG	CAG	GGT	CCT	CG	ATG	ATG	TTC	CAG	TT		960	
Qy	321	Pro	Val	Pro	Ala	Asn	Gln	Ser	Phe	Trp	Thr	Arg	Glu	Glu	Leu	Tyr	Ser	Asn	Leu	Lys	Leu	Pro			340	
Db	961	CCT	GT	ACC	AGC	TAA	CAG	TCA	TAT	TTG	GAC	TCC	GAG	AGG	ATAT	TTC	TAA	CT	TAA	AGC	TCCC				1020	
Qy	341	Arg	Ala	Ser	His	Lys	Ala	Val	Val	Asn	Gly	Asn	Leu	Met	Trp	Val	Val	Gly	Gly	Tyr	Met				360	
Db	1021	AG	AGC	ATC	TC	TAA	AGC	TGT	GGT	CA	TG	GAA	CA	TAT	GTG	GGT	TGT	TG	GAG	GAT	ATG				1080	
Qy	361	Phe	Asn	His	Ser	Asp	Tyr	Asn	Met	Val	Leu	Ala	Tyr	Asp	Leu	Ala	Ser	Arg	Glu	Trp	Leu				380	
Db	1081	TT	CA	ACC	CA	CTC	AG	AT	TAA	CA	TG	GGT	TCT	AG	CGT	ATG	ACC	TTC	CTT	CT	AGG	AGT	GGC	TT	1140	
Qy	381	Pro	Leu	Asn	Arg	Ser	Val	Asn	Asn	Val	Val	Val	Arg	Tyr	Gly	His	Ser	Leu	Ala	Leu	Tyr				400	
Db	1141	CC	ACT	AA	CCG	TCT	GTG	AA	CA	ATG	TGT	TTG	TAG	AT	ATG	TG	CT	CAT	CTT	TGG	CA	TAT			1200	
Qy	401	Lys	Asp	Lys	Ile	Tyr	Met	Tyr	Gly	Lys	Ile	Asp	Pro	Thr	Gly	Asn	Val	Thr	Asn	Glu					420	
Db	1201	AAG	GAT	AAA	AT	TAC	ATG	TAT	TGG	AGG	AAA	AT	TG	AT	CCA	CTG	GGA	ATG	TG	CA	CA	AT			1260	
Qy	421	Leu	Arg	Val	Phe	His	Ile	His	Asn	Glu	Ser	Trp	Val	Leu	Leu	Thr	Pro	Lys	Ala	Lys	Glu				440	
Db	1261	TT	GAG	AGT	TTTT	TAC	ATT	TC	AT	TAG	TCA	TG	GGT	GTG	TTG	TG	ACCC	CT	TAG	GCA	AG	GAG			1320	
Qy	441	Gln	Tyr	Ala	Val	Val	Gly	His	Ser	Ala	His	Ile	Val	Thr	Leu	Lys	Asn	Gly	Arg	Val	Val				460	
Db	1321	CAG	TAT	CG	AGT	GGT	TGG	GC	ACT	CTG	CAC	ACA	TGT	TAC	ACT	GAA	GA	TGG	CCG	AGT	GGTC				1380	
Qy	461	Met	Leu	Val	Ile	Phe	Gly	His	Cys	Pro	Leu	Tyr	Gly	Tyr	Ile	Ser	Asn	Val	Gln	Glu	Tyr				480	
Db	1381	AT	GCT	GT	CT	AT	TTG	GT	CA	CT	GGC	CT	CT	AT	TG	AT	TAA	CA	CA	ATG	TC	GAG	AA	TAT	1440	
Qy	481	Asp	Leu	Asp	Lys	Asn	Thr	Trp	Ser	Ile	Leu	His	Thr	Gln	Gly	Ala	Leu	Val	Gln	Gly	Gly				500	
Db	1441	GAT	TT	GAT	TAA	GA	CA	CA	TG	GAG	TAT	TAT	TAC	AC	ACC	CAG	GGT	GGC	CTT	GT	CA	AG	GGG	GGT	1500	
Qy	501	Tyr	Gly	His	Ser	Ser	Val	Tyr	Asp	His	Arg	Thr	Arg	Ala	Leu	Tyr	Val	His	Gly	Gly	Tyr				520	
Db	1501	TAC	GC	CA	TAG	CA	TG	TTT	TAC	CA	CA	TAG	CA	CC	AGG	GGC	CC	TAT	ACG	TT	CA	GT	TG	GGT	CTAC	1560
Qy	521	Lys	Ala	Phe	Ser	Ala	Asn	Lys	Tyr	Arg	Leu	Ala	Asp	Asp	Leu	Tyr	Arg	Tyr	Asp	Val	Asp				540	
Db	1561	AAG	GC	TTT																						

Db 1921 CTGGTATTCACTCGGAACAGTGTGATGCCCATCGGAGTGAAGCCCGCTGTTTATAGCAGCA 1980  
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680  
Db 1981 GGACCTGTTATTCGGTGTGTGTGGAAACACAGGGTCTCTCAGTGTATCTCGTGGGCGCTG 2040  
Qy 681 AlaThrAspGluGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAsp 700  
Db 2041 GCAACTGATCAACAGAAGAAAAGTTAAATCAAGATGTTTTTCCAAAAGAACTCTTGAC 2100  
Qy 701 HisAspArgCysAspGlnHisThrAspCysTrpSerCysThrAlaAsnThrAsnAspCys 720  
Db 2101 CATGACAGATGTGACACAGACACAGATTTTACAGCTGTACGCCAACCAACCAATGACTGC 2160  
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
Db 2161 CACTGGTGCATACCAATGTGTGCCCAAGGATAACCCCATGTACTACTGTAAACAGAGACC 2220  
Qy 741 IlePheArgTrpGluAsnCysProLysAspAsnProMetTrpTrpCysAsnLysLysThr 760  
Db 2221 ATTTTAGGTATGAGAAATTCGCCCAAGGATAACCCCATGTACTACTGTAAACAGAGACC 2280  
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnLys 780  
Db 2281 AGCTGCAGGAGCTGTGCCCTGGACCAAGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGC 2340  
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800  
Db 2341 ATTGCCCTGCCCGAAAATCTGTGCTGATTCGTGGCATTTGTGGTAACTCATGTGTTG 2400  
Qy 801 LysIleThrThrAlaLysGluAsnTrpAspAsnAlaLysLeuPheCysArgAsnHisAsn 820  
Db 2401 ARAATTACTTCCAGAGGAGATATGACAATGCTAAATTTGTTCTGTAGGAAACACACAT 2460  
Qy 821 AlaLeuLeuAlaSerLeuThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840  
Db 2461 GCCCTTTTGGCTTCTCTTCAACCCAGAGAGGTAGAAATTTGCTCCTTAAGCAGCTGGCA 2520  
Qy 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860  
Db 2521 ATTAATCGAGTCATCTCAGAGCATGTCAGAGCTCACCTTAACCCATGGTCCGCTTCGG 2580  
Qy 861 LysIleAsnValSerTrpTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880  
Db 2581 AAGATCAATGTGTCTTACTGTCTGGAGATATGTCCTCCCATTTTACAAATAGTTTACTA 2640  
Qy 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
Db 2641 CAGTGATGCCGTCAGGCCAGTGATGCTGGATCTGTGGAATTTTATCAGAACCCAGT 2700  
Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
Db 2701 ACTCGGGAGCTGAAGCTGCAACCTGCATCAACCCACTCAATGATGTCTGTGAAGG 2760  
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
Db 2761 CCTGCAAAACACAGTCTGAAGCTGCCGAGCACCACCATGTGCCCTTGAGGACAGCATGTGGA 2820  
Qy 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960  
Db 2821 GATTGACACAGCGGAGCTCTGATGTCATGTGGTCAGCAGACATGAGCAGTGTGTGAC 2880  
Qy 961 SerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTrpTrpThrMetSer 980  
Db 2881 TCCAAATGCCTATGTGGCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTATACGATGAGC 2940  
Qy 981 ThrCysProProGluAsnCysSerGlyTrpCysThrCysSerHisCysLeuGlnPro 1000  
Db 2941 ACTGCCCCCTTGAAATTTGTTACGCTACTGTACTGTAGTCATTGCTTGAGCAACCA 3000  
Qy 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGlySer 1020  
Db 3001 GGCTGTGGCTGTACTGATCCAGCAATACTGCCAAAGGAAATGATAGAGGGTCTCC 3060

Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTrpProGlnPro 1040  
Db 3061 TATAAAGGACCAAGTGAAGATGCCCTTCGAAAGCCCTACAGGAAATTTCTATCCACAGCCC 3120  
Qy 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTrpAsnTrpSerPheIleHisCys 1060  
Db 3121 CTGCTCAATTCAGCATGTGTCTAGAGGACAGAGATACAACTGGTGGTCTTTCTTCACTGT 3180  
Qy 1061 ProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCys 1080  
Db 3181 CCAGCTGCCAATCAACGCCCAAGTAAATGCATCAATCAGAGCATCTCTGAGAGTGT 3240  
Qy 1081 GluAsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTrpGlyAspPro 1100  
Db 3241 GAGAACCTGACCAAGCAAGCACTGCAGACCTGCATATCTGGCTTCTACGGTGTATCCC 3300  
Qy 1101 ThrAsnGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThr 1120  
Db 3301 ACCAATGGAGGAAATGTGAGCCATGCATGCATGGGCGCGCTCTCTGTGCAACACC 3360  
Qy 1121 AsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCys 1140  
Db 3361 AACACGGCGCAAGTCTCTGCACCACCAAGGGCGTCAAGGGGACGAGTCCAGCTATGT 3420  
Qy 1141 GluValGluAsnArgTrpGlnGlyAsnProLeuArgGlyThrCysTrpTrpThrLeuLeu 1160  
Db 3421 GAGTAGAATAATGCATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATATCTCTCT 3480  
Qy 1161 IleAspTrpGlnPheThrPheSerLeuSerGlnGluAspAspArgTrpTrpThrAlaIle 1180  
Db 3481 ATTGACTATCAGTTCACCTTATGCTATCCAGAGATGATCGCTATTACACAGCTATC 3540  
Qy 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200  
Db 3541 AATTTTGTGCTACTCTCTGACGAACAAACAGGGATTTGGACATGTTTCATCAATGCTCC 3600  
Qy 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220  
Db 3601 AAGAAATTTCAACCTCAACATCAGCTGGCTGCCAGTTCTCAGCTGGAACCCAGGTGGA 3660  
Qy 1221 GluGluMetProValValSerLysThrAsnIleLysGluTrpLysAspSerPheSerAsn 1240  
Db 3661 GAAGAGATGCTGTGTGTTTCAAAAACCAACATTAAGAGGTACAAAGATAGTTTCTCTAAT 3720  
Qy 1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTrpValSerAsnPhe 1260  
Db 3721 GAGAAATTTGATTTTGCACCAACCAACCAATATCACATTTCTTTGTTTATGTCTAGTAATTC 3780  
Qy 1261 ThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuVal 1280  
Db 3781 ACCTGGCCCATCAAAATTCAGATTGCTCTCTCAGCAGCAACATTTTATGGACCTGGTA 3840  
Qy 1281 GlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAlaIleValVal 1300  
Db 3841 CAGTCTCTGTGATTTCTTTCAGTTGTTTCTCTCTCTTTGCTGCTGCTGCTGCTGCTGCT 3900  
Qy 1301 TrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMet 1320  
Db 3901 TGGAGATCAACAAAGATTTGTTGGCTCCAGACGTAGAGAGCACTTCTTCGAGAGATG 3960  
Qy 1321 GlnGlnMetAlaSerArgPropheAlaSerValAsnValAlaLeuGluThrAspGluGlu 1340  
Db 3961 CAACAGATGCCAGCGCTCCCTTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4000  
Qy 1341 ProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluPro 1360  
Db 4021 CCTCTCTGATCTTATTTGGGGAGTATTAAGACTGTTCCCAACCCATTCATCTGGAGCG 4080  
Qy 1361 CysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGly 1380  
Db 4081 TGTTTTGGCAACAAAGCGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4140

Qy	1381	GlyIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSer	1400
Db	4141	GGCATCCCTCCTCCTGGGCGAGTCAGGFTCTTGCTGTGGCCAGCGCCCTGGTGGACATTCT	4200
Qy	1401	GlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGln	1420
Db	4201	CAGCAGATGCGATAGTGTACAGGAGAGTACAGGACCTGTAGAAACCGAGAGCAGCAG	4260
Qy	1421	ProProAlaGlnProGlyThrCysIle	1429
Db	4261	CCCCCTGCACAGCCTGGGACCTGCATC	4287
RESULT 3			
Id	AAZ52277	standard; cDNA; 3819 BP.	
XX	AAZ52277;		
XX	18-JUL-2000	(first entry)	
XX	Human soluble attractin-2	cdNA.	
XX	Human; soluble attractin-2; immune response; macrophage; monocyte;		
KW	T cell; immunostimulant; immunosuppressed patient; cancer;		
KW	immunodeficiency syndrome; transplant; autoimmune disease; ss.		
XX	XX		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
FH	CDS	1..3819	
FT	FT	/*tag= a	
FT	FT	/product= "Soluble attractin-2"	
XX	XX		
XX	WO2000015651-A1.		
XX	23-MAR-2000.		
XX	14-SEP-1999;	99WO-US020948.	
XX	14-SEP-1998;	98US-0100137P.	
XX	(DAND )	DANA FARBER CANCER INST INC.	
XX	Duke-Cohan JS,	Schlossman SF;	
XX	WPI;	2000-271373/23.	
DR	P-PSDB;	AA70692.	
XX	Isolated nucleic acids encoding human attractin polypeptides useful for		
PT	enhancing immune responses.		
XX	Claim 3; Fig 12;	120pp; English.	
XX	The patent discloses four forms of human attractin polypeptides which		
CC	enhance immune response by promoting macrophage and monocyte spreading in		
CC	the presence of T cells. These include soluble attractin-1 and -2 and		
CC	membrane attractin-1 and -2. These various forms of attractin are encoded		
CC	by alternatively spliced mRNA molecule transcribed from a single gene.		
CC	The present sequence is a cDNA encoding human soluble attractin-2.		
CC	Attractin-2 differs from attractin-1 in having a 74 amino acid insert in		
CC	the N-terminal. This sequence can be used to enhance immune response in		
CC	immunosuppressed patients such as those undergoing chemo- and radio-		
CC	therapy treatment for cancer or those suffering from common variable		
CC	immunodeficiency syndrome. The proteins may also be used to screen		
CC	modulators (agonists and antagonists) of immune response which may also		
CC	be used to regulate immune reactions. Attractin antibodies can be used to		
CC	inhibit immune response in transplant recipients or patients afflicted		
XX	with autoimmune disease		
SQ	Sequence 3819 BP; 969 A; 899 C; 991 G; 960 T; 0 U; 0 Other;		
Alignment Scores:			
Pred. No.:	0	Length:	3819

Score:	1267.00	Matches:	1267
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	88.66%	Indels:	0
DB:	3	Gaps:	0
US-09-787-097-12 (1-1429) x AAZ52277 (1-3819)			
Qy	1	MetValAlaAlaAlaAlaAlaThrGluAlaAArgLeuAArgArgThrAlaAlaThrAla	20
Db	1	ATGTGTCCCGCAGCGCGCACTGAGGCAGGCTGAGGAGGAGGAGCGCGGCGACGCA	60
Qy	21	AlaLeuAlaGlyAArgSerGlyGlyProHisTrpAspTrpAspValThrAArgAlaGlyAArg	40
Db	61	GCAGTCGCGGCGAGGACGCGCGGCGGCATCGGACTGGAGCTGGAGCTGACAGGGCTGGAGG	120
Qy	41	ProGlyLeuGlyAlaGlyLeuAArgLeuProAArgLeuLeuSerProProLeuAArgProAArg	60
Db	121	CCGCGGCTGGCGGCGCGGCTCGCGCTCCCGCGGCTGTGTCTCCACGCTGCGCGCACGG	180
Qy	61	LeuLeuLeuLeuLeuLeuLeuProProProLeuLeuLeuLeuLeuLeuProCysGlu	80
Db	181	CTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT	240
Qy	81	AlaGluAlaAlaAlaAlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGlu	100
Db	241	GCCGAGCGCGCGCGCGCGCGGCGGTGTGGGCTCAGCCGAGCCGAGGCGCAAGGAA	300
Qy	101	CysAspAArgProCysValaAnGlyGlyAArgCysAsnProGlyThrGlyGlnCysValCys	120
Db	301	TGTACCGGCGCTGTGTCAACGCGGCTGTGTCAACCCCTGGCACCAGGCGGCTGTGTGTGT	360
Qy	121	ProAlaGlyTrpValGlyGluGlnCysGlnHisCysGlyGlyAArgPheAArgLeuThrGly	140
Db	361	CCGCGCGCTGGGTGGCGAGCAATGCCAGCATGTGGGGGCGCTTCAGACTAACTGGA	420
Qy	141	SerSerGlyPheValThrAspGlyProGlyAsnTrpLysTrpLysCysThrTrp	160
Db	421	TCCTCTGGGTTTGTGACAGATGGACCTTGGAAATATATAATACAAAAGAGTGCACGTGG	480
Qy	161	LeuIleGluGlnProAsnAArgIleMetAArgLeuAArgPheAsnHisPheAlaThrGlu	180
Db	481	CTCATGGAAGACAGCCAAATAGAAATAGAGACTTCGTTTCAATCATCTTTTGTACAGAG	540
Qy	181	CysSerTrpAspHisLeuTrpValTrpAspGlyAspSerIleTrpAlaProLeuValAla	200
Db	541	TGTAGTTGGACCATTTATATGTTATATGATGGGACTCAATTTATGACCGCTAGTTGCT	600
Qy	201	AlaPheSerGlyLeuIleValProGluAArgAspGlyAsnGluThrValProGluValVal	220
Db	601	GCATTTAGTGGCTCATGTGTCTCGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTT	660
Qy	221	AlaThrSerGlyTrpAlaLeuLeuHisPheSerAspAlaAlaTrpAsnLeuThrGly	240
Db	661	GCCACATCAGGTTATGCTTGTCTGCAATTTTGTAGTGATGTCTTATATTTTACTGCTGA	720
Qy	241	PheAsnIleThrTrpSerPheAspMetCysProAsnAsnCysSerGlyAArgGlyCys	260
Db	721	TTTAAATATTACTTACAGTTTTGATGTGTCCAAATAACTGTCTCAGGCCGAGGAGAGTGT	780
Qy	261	LysIleSerAsnSerSerGluThrValGluCysGluCysSerGluAsnTrpLysGlyGlu	280
Db	781	AGATCAGTAAATAGCAGCGAAACTGTTGATGTGAATGTCTGAAAACTGGAAAGGTGAA	840
Qy	281	AlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisAArgGlyIleCysAsn	300
Db	841	GCATGTGCATCTCTCACTGTACAGACAACTGTGGTTTCTCTCATCAGGCGCATCTGCAAT	900
Qy	301	SerSerAspValAArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerVal	320
Db	901	TCAAGTGATGTACAGAGATGCTCTGCTCTTCAGACTGGCAGGCTCTGATGTTTCAGTT	960
Qy	321	ProValProAlaAsnGlnSerPheTrpThrAArgGluTrpSerAsnLeuLysLeuPro	340

Alignment Scores:  
Pred. No.:

Db 961 CCTGTACCAGCTAACGAGTCATTTTGGACTCGAGAGAAATATCTTAACCTTAAGCTCCCC 1020  
Qy 341 ArgAlaSerHisLysAlaValValAlaAsnGlyAsnIleMetTrpValValGlyGlyTyrMet 360  
Db 1021 AGAGCATCTCATAAAGCTGTGGTCAATGGGAAACATTTATGTGGGTTTGTGGAGATATATG 1080  
Qy 361 PheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeu 380  
Db 1081 TTCAACCACTCAGATTATACATGCTGTAGCGGTATGACCTTGCTTCTAGGAGTGGCTT 1140  
Qy 381 ProLeuAsnArgSerValAsnAsnValValAlaArgTyrGlyHisSerLeuAlaLeuTyr 400  
Db 1141 CCACCTAAACCGTTCTGTGAACATGTGGTTGTTAGATATGGTTCATTTCTTGGCATATAC 1200  
Qy 401 LysAspLysIleTyrMetTyrGlyGlyLysIleAspProThrGlyAsnValThrAsnGlu 420  
Db 1201 AAGGATAAAATTTACATGTATGAGGAGAAATTTGATCCAACTGGGAATGTGACCAATGAG 1260  
Qy 421 LeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGlu 440  
Db 1261 TTGAGAGTTTTCACATTTCAATAGTGTATGGGTGCTTGTGACCCCTTAAGGCAGAGAG 1320  
Qy 441 GlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValVal 460  
Db 1321 CAGTATGCACTGGTGGGCACTCTGCACACATTTGTACACTGAAGAATGGCGAGTGGTC 1380  
Qy 461 MetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyr 480  
Db 1381 ATGCTGGTCACTTTGGTCTACTGCCCTCTCTATGGATATATTAAGCATGTGAGGANTAT 1440  
Qy 481 AspLeuAspLysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGly 500  
Db 1441 GATTTGGATAAAGAACATGAGTATATTACACACCGAGGTGCCCTTGTGCAAGGGGT 1500  
Qy 501 TyrGlyHisSerSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyr 520  
Db 1501 TAGGGCCATAGCAGTGTTCACGACCATAGGACCGAGGCGCCCTATACGTTTATGGTGGCTAC 1560  
Qy 521 LysAlaPheSerAlaAsnLysTyrArgLeuAlaAspLeuTyrArgTyrAspValAsp 540  
Db 1561 AAGGCTTTCAGTCCCAATAGTACCGCTTGCCAGATGATCTCTACCGATATGATGGAT 1620  
Qy 541 ThrGlnMetTrpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaVal 560  
Db 1621 ACCCAGATGTGGACCACTTCTTAAGGACGCGGATTTTCCGTTACTGTGCACACAGCTGTG 1680  
Qy 561 IleValSerGlyThrMetLeuValPheGlyGlyValAsnThrHisAsnAspThrSerMetSer 580  
Db 1681 ATAGTGTAGTGAACCAATGCTGGTGTGTTGGGGAAACACACACATGACATCTATGAGC 1740  
Qy 581 HisGlyAlaLysCysPheSerAspPheMetAlaTyrAspIleAlaCysAspArgTrp 600  
Db 1741 CATGGCGCAAAATGCTTCTTCCAGATTTCATGGCCTATGACATTCGCTGTGACCGCTGG 1800  
Qy 601 SerValLeuProArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaVal 620  
Db 1801 TCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATTGGCCATTTGAGCAGTTC 1860  
Qy 621 LeuHisAsnSerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIle 640  
Db 1861 TTACACACAGCACCATGATGTGTTGGTGGTGTTCATAGTCTCCTCCTCAGCGACATC 1920  
Qy 641 LeuValPheThrSerGlnGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAla 660  
Db 1921 CTGGTATTCACCTCGGAACAGTGTGAGCGCATCGGAGTGAAGCCCTTGTATTAGCAGCA 1980  
Qy 661 GlyProGlyIleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeu 680  
Db 1981 GGACCTGGTATTCGGTGTGTGCGAACACACAGAGGTGCTCTCAGTGTATCTCGTGGCGCTG 2040  
Qy 681 AlaThrAspGlnGlnGlnGlyLeuLysSerGluCysPheSerLysArgThrLeuAsp 700

Db 2041 GCAACTGATGAACAAGAAAGAAAGTTAAAAATCAGAATGTTTTTCCAAAAGAACTCTTGAC 2100  
Qy 701 HisAspArgCysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCys 720  
Db 2101 CATCACAGATGTGTACCAACAGCACACAGATTGTTCAGCTGTACAGCCAAACCAATGACTGC 2160  
Qy 721 HisTrpCysAsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSer 740  
Db 2161 CACTGGTGCATGACCATTTGTGTCCCGAGAACACACAGCTGCTCAGAAAGCCAGATCTCC 2220  
Qy 741 IlePheArgTyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThr 760  
Db 2221 ATTTTTAGGTATGAAATTGCCCAAGATAACCCCATGACTACTACTGTAAACAAGAGACC 2280  
Qy 761 SerCysArgSerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCys 780  
Db 2281 AGCTGCAGGAGCTGTGCTCTGGACCAAGACTGCCAGTGGAGCCCGGANTCAGAGTGC 2340  
Qy 781 IleAlaLeuProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeu 800  
Db 2341 ATTGCCCTGCCCGAAAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAATCTCATGTTT 2400  
Qy 801 LysIleThrThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsn 820  
Db 2401 AAAATTACTACTGCCAAGGAGAAATTATGCAATGCTTAATTTGTTCTGTAGAACCAACAT 2460  
Qy 821 AlaLeuLeuAlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArg 840  
Db 2461 GCCCTTTTGGCTCTCTTACACCCAGAGAGTAGAATTTGCTCTTAAGCAGCTGCGA 2520  
Qy 841 IleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArg 860  
Db 2521 ATAAATCAGATCATCTCAGACATGTCAGACTCACCTTAACCCCATGGGTGGCCCTTCGG 2580  
Qy 861 LysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeu 880  
Db 2581 AAGATCAATGTGCTCTACTGTCTGGAGATATGTCCTCCCAATTTACAAATAGTTACTA 2640  
Qy 881 GlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSer 900  
Db 2641 CAGTGGATGCTGTGAGCCAGCTGATGCTGGATTTCTGTGCAATTTTATCAGAACCCAGT 2700  
Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
Db 2701 ACTCGGAGCTGAAGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGG 2760  
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
Db 2761 CCTGCAAAACACACAGTGTCTAAGCAGTGGCGACACCATGTGCTTGGAGCAGCATGTGA 2820  
Qy 941 AspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960  
Db 2821 GATTGCAACAGCGGAGCTCTGAGTGCATGTGTGGTCAGCAACATGAAGCAGTGTGGAC 2880  
Qy 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSer 980  
Db 2881 TCCCAATGCCATGTGGCTCTCTCTCTTGGCCAGTGTATGGAATGGTATACGATGAGC 2940  
Qy 981 ThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnPro 1000  
Db 2941 ACCTGCCCTCCCTGAAAAATTTGTCAGGCTACTGTACTCTGTAGTGTATGCTTTGGAGCA 3000  
Qy 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020  
Db 3001 GGCTGTGGCTGGTGTACTGTATCCAGCAATACTGGCAAGGAAATGATAGAGGGTTC 3060  
Qy 1021 TyrLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnPro 1040  
Db 3061 TATAAAGGACCGATGAAGATGCCCTTCGCAAGCCCTTACAGAAATTTCTATCCACAGCCC 3120  
Qy 1041 LeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCys 1060  
Db 3121 CTGCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTCTTCACTCTGT 3180



```
QY 1061 ProAlaCysGlnCysAsnGlyHisSerIysCysIleAsnGlnSerIleCysGlnIysCys 1080
DB 3181 CCAGCTTGCCAAATGCAACGGCCACAGTAATGATCATCAATCAGAGCATCTGTGAGAGGTGT 3240
QY 1081 GluAsnLeuThrThrGlyIysHisCysGluThrCysIleSerGlyPheTyrGlyAspPro 1100
DB 3241 GAGACCTGACACACAGGCAAGCACTGCGAGACTGCATATCTGGGTTCTACGGTGTATCC 3300
QY 1101 ThrAsnGlyGlyIysCysGlnProCysIysCysAsnGlyHisAlaSerLeuCysAsnThr 1120
DB 3301 ACCAATGAGGGAATGTGAGCATGTCAAGTCAATGGCGCATCTCTGTGCAACACC 3360
QY 1121 AsnThrGlyIysCysPheCysThrThrIysGlyValIysGlyAspGluCysGlnLeuCys 1140
DB 3361 AACCGGGCAAGTGTCTTCACACCAAGGGCGTCAAGGGGAGCAGTGCAGGTATGT 3420
QY 1141 GluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeu 1160
DB 3421 GAGGTAGAAATCGATACCAAGGAACCCCTCTCAGAGGAACATGTTATTACTCTCTT 3480
QY 1161 IleAspTyrGlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIle 1180
DB 3481 ATTGACTATCAGTTCACCTTTAGTCTATCCAGGAAGATGATCGGTATTACACAGCTATC 3540
QY 1181 AsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSer 1200
DB 3541 AATTTGTGCTACTCTCAGCAACAAACAGGGATTTGGACATGTTCAATCAATGCCCTCC 3600
QY 1201 LysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGly 1220
DB 3601 AAGAATTTCAACCTCAACATCACCTGGGCTGCAGTTTCTCAGCTGGAACCCAGGCTGGA 3660
QY 1221 GluGluMetProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsn 1240
DB 3661 GAAGAGATCCCTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATGTTCTCTAAT 3720
QY 1241 GluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPhe 1260
DB 3721 GAGAGTTTGTATTTTCGCAACCAACCAATATACATTTCTTTGTTTATGTCTAGTAAATTC 3780
QY 1261 ThrTrpProIleIysIleGln 1267
DB 3781 ACCTGGCCCATCAAAATTCAG 3801

RESULT 4
AAZ91920
ID AAZ91920 standard; cDNA; 8590 BP.
XX AAZ91920;
AC AAZ91920;
XX
XX 08-JUN-2000 (first entry)
XX
XX Human mahogany protein coding sequence #2.
DE
XX
XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
KW weight regulation; cell therapy; body weight disorder; cachexia;
KW anorexia; hyperpigmentation; increased metabolic rate disorder;
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200005373-A2.
XX
XX 03-FEB-2000.
XX
XX 21-JUL-1999; 99WO-US016484.
XX
XX 21-JUL-1998; 98US-0093630P.
XX
XX 20-OCT-1998; 98US-0104578P.
XX
XX 05-FEB-1999; 99US-00245041.
XX
XX (MILL-) MILLENIUM PHARM INC.
PA
```

```
XX Moore K, Nagle DL;
PI
XX WP; 2000-195103/17.
DR P-PSDB; AAY81807.
XX
XX New human and murine mahogany genes, useful, e.g. for diagnosis and
PT treatment of body weight disorders.
PT
XX Claim 1; Fig 18a; 188pp; English.
PS
XX This sequence represents a human mahogany gene of the invention. The
XX mahogany genes are used: (i) to produce recombinant mahogany (mg)
XX proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming
XX therapeutics; (iii) as a source of diagnostic probes and primers for
XX detecting expression of mg genes or mutations; regulatory defects, in
XX this gene, or for isolation of related sequences; and (iv) in (cell-
XX based) gene therapy. (ii) are used to raise specific antibodies (Ab); to
XX identify other (extra)cellular products involved in weight regulation,
XX and to screen for agents that disrupt interaction between (II) and other
XX macromolecules. The Ab are used to detect abnormal levels (or function)
XX of (II) for diagnosis, prognosis or monitoring of treatment); to
XX evaluate (II)-expressing cells intended for cell therapy, and as
XX therapeutic mg inhibitors. Cells that express the mg gene, (or contain the
XX mg polypeptide) are used to identify agents (A) that modulate mg
XX activity. (A) are potentially useful for the treatment of body weight
XX disorders, particularly obesity, cachexia or anorexia, or other
XX conditions associated with the mg gene such as hyperpigmentation,
XX hyperphagia and disorders that result in increased metabolic rate
XX
XX Sequence 8590 BP; 2275 A; 1990 C; 2061 G; 2264 T; 0 U; 0 Other;
SQ
XX
XX Alignment Scores:
XX Pred. No.: 0 Length: 8590
XX Score: 1141.00 Matches: 1341
XX Percent Similarity: 99.70% Conservative: 0
XX Best Local Similarity: 99.70% Mismatches: 2
XX Query Match: 79.85% Indels: 4
XX DB: 3 Gaps: 0
XX
XX US-09-787-097-12 (1-1429) x AAZ91920 (1-8590)
QY 87 AlalaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
DB 256 GCGCGCGGTGTCCGGCTCGCGCGCAGCCGAGCCCAAGGAATGTGACCGCGCGGTGTGTC 315
QY 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTyrValGly 126
DB 316 AACGCGGTGCTGCAACCTGGCACCGCGCAGTGTCTGCCCCCGCGCGGTGGCGC 375
QY 127 GluGlnCysGlnHisCysGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146
DB 376 GAGCAATGCCAGCACTCGGGGCGCGCTTCAGACTAATCGATCTCTGGGTTTGTGACA 435
QY 147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166
DB 436 GATGGACCTGGAATTAATAACAAACGAAGTGCACGTGGCTCATTAAGAGCACACCA 495
QY 167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186
DB 496 AATAGAAATATGAGACTTCGTTTCAATCAATTTTGTCTACAGAGTGTAGTGGACCATTA 555
QY 187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
DB 556 TATGTTTATGATGGGACTCAATTTATGCACCGGTAGTGTGTCATTTAGTGGCTCAT 615
QY 207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226
DB 616 GTTCTCTGAGAGATGGCAATGAGACTGTCCCTGAGGTTGTGCCACATCAGGTATGCC 675
QY 227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrThrSer 246
DB 676 TTGCTGCATTTTTTTTAGTGATGCTGTATAAATTTGACTGGATTATATATTACTACAGT 735
```



QY 247 PheAspMetCysProAsnIleAsnValArgGlyGluCysIleSerAsnSerSer 266  
Db 736 TTTGATATGTGTCACAAATTAACCTCTCAGCCGAGGAGGTGAAGATCAGTAAATAGCAGC 795  
QY 267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286  
Db 796 CA-TACTGTGATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCA 854  
QY 286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgGl 306  
Db 855 CTGTACAGACAACATGTGGTTCCTCATCGAGGCATCTGCAATTCAGTGAATGTGAGAG 914  
QY 306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnGl 326  
Db 915 ATGCTCTGCTCTCAGACTGGCAGGCTCTGATGTTGAGTTCCTGTACAGCTTAACCA 974  
QY 326 nSerPheTrpThrArgGluGluTrpSerAsnLeuLysLeuProArgAlaSerHisLysAl 346  
Db 975 GTCAATTTGGACTCGAGAGGAATTTCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGC 1034  
QY 346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTrpMetPheAsnHisSerAspTy 366  
Db 1035 TGTGGTCAATGGAAACATATGTGGGTGTGGAGGATATATGTTCAACCATCTCAGATTA 1094  
QY 366 rAsnMetValLeuAlaTrpAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386  
Db 1095 TAACTAGGTCTAGCTATGACCTTGTCTTCTAGGAGTGGCTTCCACTAAACCGTTCGT 1154  
QY 386 lAsnAsnValValArgTrpGlyHisSerLeuAlaLeuTrpLysAspLysIleTrpMe 406  
Db 1155 GAACAATGTGTTGTAGATATGTCATCTTTGGCATTTATACAGGATATAAATTTACAT 1214  
QY 406 tTrpGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426  
Db 1215 GTATGAGGAAAAATTTGATTC-ACTGGGATGTGACCAATGATGTTGAGAGTTTTCACA 1273  
QY 426 lHisAsnGluSerTrpValLeuLeuTrpProLysAlaLysGluGlnTrpAlaValValG 446  
Db 1274 TTCATAATGAGTCATGGTGTGTCACCCCTAAGGCAAGGAGCAGTATGACGTGGTGTG 1333  
QY 446 lHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466  
Db 1334 GGCACTCTGCACACATTTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCATCTTG 1393  
QY 466 lHisCysProLeuTrpGlyTrpIleSerAsnValGlnGluTrpAspLeuAspLysAsnT 486  
Db 1394 GTCACTGCCCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTTGGATAGAACA 1453  
QY 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTrpGlyHisSerSerV 506  
Db 1454 CATGGAGTATATTACACACCCAGGGTGGCCCTGTGCAAGGGGTACGGCCATAGCAGTG 1513  
QY 506 alTrpAspHisArgThrArgAlaLeuTrpValHisGlyTrpLysAlaPheSerAlaA 526  
Db 1514 TTTACGACCATAGACACAGCCCTATACGTTATGTTGGCTACAAGGCTTTTCAGTGCCA 1573  
QY 526 snLysTrpArgLeuAlaAspAspLeuTrpArgTrpAspValAspThrGlnMetTrpThrI 546  
Db 1574 ATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGAGATACCCAGATGTGGACCA 1633  
QY 546 lLeuLysAspSerArgPheArgTrpLeuHisThrAlaValIleValSerGlyThrM 566  
Db 1634 TTTTAAAGGACACCGCAATTTTCGTTACTTGGACACAGCTGTGATGAGTGGAGNACCA 1693  
QY 566 etLeuValPheGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586  
Db 1694 TGCTGGTGTGGGGGAAACACACAATGACACATCTATGACCATGAGCCATGGCGCAATGCT 1753  
QY 586 heSerSerAspPheMetAlaTrpAspIleAlaCysAspArgTrpSerValLeuProAtgP 606  
Db 1754 TCTCTTCAGATTTTCATGGCCCTATGACATTCCTGTGACCGCTGGTGTGCTTCCAGAC 1813

QY 606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626  
Db 1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGCTTTACACAAACAGACCA 1873  
QY 626 etTrpValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerG 646  
Db 1874 TGTATGTGTTGGTGGTTCATATAGTCTCTCTCCTCAGCAGCATCTCTGGTATTCACCTCG 1933  
QY 646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
Db 1934 AACAGTGTGATGCGCATCGGAGTGAAGCCGCTGTTTGTAGCAGCAGGACCTGGTATTCGGT 1993  
QY 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686  
Db 1994 GTGTGTGGACACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAG 2053  
QY 686 luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706  
Db 2054 AAGAAAAAGTTTAAAAATCAGATGTTTTTCCAAAAAGAACTCTTGACCATGACAGATGTGACC 2113  
QY 706 lnHisThrAspCysTrpSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726  
Db 2114 AGCACACAGATTTTACAGCTGCACAGCCAAACCAATGACTGCCACTGTGTGCAATGACC 2173  
QY 726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTrpGluA 746  
Db 2174 ATTGTGTCCCAGGAACACACAGCTGTCTCAGAAGCCAGATCTCCATTTTTAGTATGAGA 2233  
QY 746 snCysProLysAspAsnProMetTrpTrpCysAsnLysLysThrSerCysArgSerCysA 766  
Db 2234 ATTGCCCAAGGATTAACCTTATGCTACTGTACAGAGACAGCAGCTGCAGAGCTGTG 2293  
QY 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786  
Db 2294 CCCTGGACCAAGAACTGCCAGTGGGAGCCCGGAATCAGGAGTGCATTTGCCCTGCCGAAA 2353  
QY 786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrAlaAl 806  
Db 2354 ATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAACTCATGTTTGAATAATTTACTACTGCCA 2413  
QY 806 ysGluAsnTrpAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826  
Db 2414 AGGAGATTTATGACAAATGCTAAATTTGTTCTGTAGGAACCAACAATGCCCTTTGGCTTCT 2473  
QY 826 euThrThrGlnLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846  
Db 2474 TTCAACCCAGAGAAGGTAGAATTTGCTCTTAAGCAGCTGCGAATAATCGACTCATCTC 2533  
QY 846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866  
Db 2534 AGACATGTCCAAGTCACTTACCCCATGGGTGGCTTCGGAAGATCAATGTGTCTCT 2593  
QY 866 yTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886  
Db 2594 ACTGGTGTGGGAAGATATGTCCCAATTTACAAATAGTTTTACTACAGTGGATGCCGCTG 2653  
QY 886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906  
Db 2654 AGCCAGTGTGCTGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAGG 2713  
QY 906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926  
Db 2714 CTGCAACCTGTCATCAACCCACTCAATGGTAGTGTCTGTGAAGCCCTGCAACACACAGTG 2773  
QY 926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946  
Db 2774 CTAGCAGTGGCGGACACCATGTGCTGTGAGACAGCATGTGAGATTTGCCAGCGGCA 2833  
QY 946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTrpValA 966  
Db 2834 GCTCTGAGTGCATGTGTGTCAGCAACATGAAGCAGTGTGTGGACTCCAAATGCTATGTGG 2893  
QY 966 laSerPhePropheGlyGlnCysMetGluTrpTrpThrMetSerThrCysProGluA 986



CC cancer or those suffering from common variable immunodeficiency syndrome.  
 CC The proteins may also be used to screen modulators (agonists and  
 CC antagonists) of immune response which may also be used to regulate immune  
 CC reactions. Attraction antibodies can be used to inhibit immune response in  
 CC transplant recipients or patients afflicted with autoimmune disease  
 XX  
 SQ Sequence 4068 BP; 1042 A; 948 C; 1027 G; 1051 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 0 Length: 4068  
 Score: 1123.00 Matches: 1323  
 Percent Similarity: 99.70% Conservative: 0  
 Best Local Similarity: 99.70% Mismatches: 2  
 Query Match: 78.59% Indels: 4  
 DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x AAZ52275 (1-4068)

QY 105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124  
 DB 91 TGTGTCAACGGCGTGTGTCGAACCTGGCACCAGCGGCGTGTGCGCCCGCGGCTGG 150  
 QY 125 ValGlyGluClnCysGlnHisCysGlyArgPheArgLeuThrGlySerSerGlyPhe 144  
 DB 151 GTGGGGGAGCAATGCCAGACTCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTT 210  
 QY 145 ValThrAspGlyProGlyAsnTyrLysThrLysCysThrTrpLeuIleGluGly 164  
 DB 211 GTGACAGTGGACCTGGAAATTATAAATACAAACGAGTGCAGCGGCTCATTTGAGGA 270  
 QY 165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184  
 DB 271 CAGCCAAATAGAATAATGACACTCGTTTCAATCATTTTGCTACAGAGTGTAGTTGGGAC 330  
 QY 185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGly 204  
 DB 331 CATTTATATGTTTATGATGGGACTCAATTTATGACACCGCTAGTGTCTGCTATTAAT 390  
 QY 205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGly 224  
 DB 391 CTCATTTGTTCTCAGAGAGATGCAATGAGACTGTCCTGAGGTTGTTGCCACATCAGGT 450  
 QY 225 TyrAlaLeuLeuHisPheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThr 244  
 DB 451 TATGCTTGTGTCATTTTTTATGATGCTGCTTAAATTTGACTGGATTTAATATTACT 510  
 QY 245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264  
 DB 511 TACAGTTTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGGTGAAGATCAGTAAT 570  
 QY 265 SerSerGlu-ThrValGluCysGlnCysSerGluAsnTrpLysGlyGluAlaCysAspIle 284  
 DB 571 AGCAGGCA-TACTGTTCAATGTCAATGTTCTGAAACTGGAAGGTGAAGCATGTGACAT 629  
 QY 284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304  
 DB 630 TCCTCACTGTACAGACACACTGTGTTTCTCATCGAGCACTGCAATTCAGTGATGT 689  
 QY 304 LArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAl 324  
 DB 690 CAGAGGATGTCCTGCTTCTCAGACTGGCAGGGTCTCGATGTTTCAGTTCTCTGTACCCAG 749  
 QY 324 aAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHi 344  
 DB 750 TAACCACTCAATTTTGGACTCGAGAGGAATATCTTAACCTTAAGCTCCCGAGGACATCTCA 809  
 QY 344 sLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSe 364  
 DB 810 TAAAGCTGTGTTCAATGGAAACATTATGTGGTGTGGTGTGGAGGATATATGTTCAACCACTC 869  
 QY 364 rAspTyrAsnMetValIleAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAr 384  
 DB 870 AGATTATAACATGTTCTAGCGTATGACCTTGCTTCTAGGGAGTGGCTTCCACTTAACCG 929

QY 384 qSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIle 404  
 DB 930 TTCGTGACCAATGGTGTGTAGATATGTCATCTTTGGCATATATACAGGATAAAT 989  
 QY 404 eTyrMetTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424  
 DB 990 TTACATGTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTGTGAGATTT 1048  
 QY 424 heHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluTyrAlav 444  
 DB 1049 TTCACATTCAATAGTGTGATGGGTGTGTGACCCCTAAGGCAAGAGGACATGATGTCAG 1108  
 QY 444 alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464  
 DB 1109 TGGTTGGGCACTCTGCACACATTTGTACACTGAAGATGCCCGAGTGTGTCATGCTGTCA 1168  
 QY 464 lePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspL 484  
 DB 1169 TCTTTGGTCACTGCGCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATA 1228  
 QY 484 ysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHis 504  
 DB 1229 AGACACATGAGATATATTACACCCAGGTCCTTGTGCAAGGGGTACGGCCATA 1288  
 QY 504 erSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS 524  
 DB 1289 GCAGTGTATTACGACCATAGGACAGGCGCTTATACGTTCTATGTTGCTACAGGCTTTCA 1348  
 QY 524 erAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMet 544  
 DB 1349 GTGCCAATAAGTACCGCTTGCAGATGATCTTACCGATATGATGTGGATACCCAGATG 1408  
 QY 544 rpThrIleLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerG 564  
 DB 1409 GGACCATCTTAAAGACAGCGGATTTTTCGTTTAC-TGACACACAGCTGTGATAGTGAGTG 1468  
 QY 564 lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584  
 DB 1469 GAACCATGCTGGTGTGTTGGGGGAAACACACAAATGACACATCTATGAGCCATGGGCGCA 1528  
 QY 584 ysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTyrSerValLeuP 604  
 DB 1529 AATGCTTCTCTCAGATTTTATGGCTATGACATTCCTGTGACCGTGTGATGCTTC 1588  
 QY 604 roArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnS 624  
 DB 1589 CCAGACCTGATCTCCACCATGATGTCAACAGATTTGGCCATTCAGCAGTCTTACACAA 1648  
 QY 624 erThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheP 644  
 DB 1649 GCACATGATATGTTGCTGGTGTGTTCAATAGTCTCTCTCAGCAGCATCTCTGATTTCA 1708  
 QY 644 hrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyI 664  
 DB 1709 CCTCGAAACAGTGTGATGCGCATCGGAGTGAAGCGCTTGTAGCAGCAGGACCTGGTA 1768  
 QY 664 leArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspG 684  
 DB 1769 TTCGGTGTGTGGAAACACAGGGTGTGCTCAGTGTATCTCGTGGGCTGGCACTGATG 1828  
 QY 684 luGlnGluGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgC 704  
 DB 1829 AACAAAGAGAAAAGTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGAT 1888  
 QY 704 ysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysA 724  
 DB 1889 GTGACACAGCACACAGATTTGTACAGCTGCACGCCAACCAATGATGTCCTGCTGGTGA 1948  
 QY 724 snAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgT 744  
 DB 1949 ATGACCATTTGTGTCCTCCAGGACACACAGCTGCTCAGAGGCCAGATCTCTCCATTTTAGGT 2008

744 yrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgS 764  
Db ATGAGAAATGCCCCAAGGATACCCCTATGTAATCTGTAACAGAGAACCCAGCTGAGGA 2068  
764 erCysAlaLeuAspGlnAsnCysGlnTyrProArgAsnGlnGluCysIleAlaLeuP 784  
Db GCTGTGCCCTGGACCAAGAACTCCAGTGGAGGCCCGGAATCAGGAGTGCATTGGCCCTGC 2128  
784 roGluAsnIleCysGlyIleGlyTyrHisLeuValGlyAsnSerCysLeuLysIleThrT 804  
Db CCGAAATATCTGTGCAATGCTGGCAATGCTGGAAATCAATGTTGAAATACTA 2188  
804 hrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuA 824  
Db CTGCCAAGGAGAATATGCAATGCTAAATGCTTCTGTAGGAACCAACCAATGCCCTTTGG 2248  
824 laSerLeuThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnS 844  
Db CTCTCTTCAACCCCAAGAGAGGTAGAAATGCTTAAAGCAGCTGGCAATGAAATGCAAT 2308  
844 erSerGlnSerMetSerLysLeuThrLeuThrProTyrValGlyLeuArgLysIleAsnV 864  
Db CATCTCAGAGCATGCTCAAGCTCACCTTAACCCCATGGCTGGCTTGGAGATCAATG 2368  
864 alSerTyrTyrCysTyrGluAspMetSerProPheThrAsnSerLeuLeuGlnTyrMetP 884  
Db GTCCCTACTGGTCTGGGAGATATGTCCCAATTTACAATATGTTTACTACAGTGGATGC 2428  
884 roSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyL 904  
Db CGTCTGAGCCCAAGTGTGTAATCTGTGGAAATTTATCAGAACCCAGTACTCGGGAC 2488  
904 euLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnH 924  
Db TGAAGGCTGCAACTGATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCCTGCAAAAC 2548  
924 isSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrS 944  
Db ACAGTCTAAGCAGTCCCGGACACCATGTGCCCTTGGAGACAGCATGTGGAGATTGCACCA 2608  
944 erGlySerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAspSerAsnAlaT 964  
Db GCGGAGCTCTGAGTGCATGTGGTGAGCAACATGAAGCAGTGTGTGGACTCCAATGCT 2668  
964 yrValAlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSerThrCysProp 984  
Db ATGTGGCCCTTCCCTTGTGCCCAGTGTATGGAATGTGTATACGATGAGCACCTGCCCCC 2728  
984 roGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyT 1004  
Db CTGAAATTTGTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCT 2788  
1004 rpCysThrAspProSerAsnThrGlyLysGlyCysIleGluGlySerTyrLysGlyP 1024  
Db GGTGTACTGATCCAGCAATACTGGCAAGGAAATGCATAGAGGTTCTCTATAAGGAC 2848  
1024 roValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnS 1044  
Db CAGTGAAGATGCTTCGCAAGCCCTTACAGGAATTTCTATCCAGCCCCCTGCTCAAT 2908  
1044 erSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHisCysProAlaCysG 1064  
Db CCAGCATGTGTCTAGAGGACAGCAGATACAATGTGTCTTTTCTTCTTCTTCTTCTTCTT 2968  
1064 lnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGlnLysCysGluAsnLeuT 1084  
Db AATGCAACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAATGTGAGAACCTGA 3028  
1084 hrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyV 1104  
Db CCACAGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAATGAG 3088  
1104 lLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyL 1124

Db GGAATGTGACCATGCAAGTGGCAATGGCAGCGCTCTCTGTGCAACACCAACACGCGCA 3148  
QY ysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluA 1124  
Db AGTCTCTTCTGCACCAAGGGCGTCAAGGGGACAGTGCAGCTATGTGAGGTAGAAA 3208  
QY snArgTyrGlnGlnValAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrG 1164  
Db ATCGATACCAAGGAACCCCTCTCAGAGAAACATGTTATTATATCTCTTCTTATTGACTATC 3268  
QY lnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheValA 1184  
Db AGTTCACTTCTAGTCTATCCAGGAGATGCTATTACACAGCTATCAATTTTGTGG 3328  
QY laThrProAspGluGlnAsnArgAspMetPheIleAsnAlaSerLysAsnPheA 1204  
Db CTACTCTGAGAACAAACAGGATTTGGACATGTTCAATATGCTCTCCAGAAATTTCA 3388  
QY snLeuAsnIleThrTyrAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetP 1224  
Db ACCTCAACATCACCTGGCTGCCAGTTTCTCAGCTGGAACCCAGGCTGGAGAGAGATGC 3448  
QY roValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheA 1244  
Db CTGTGTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTCTAAATGAGAAGTTG 3508  
QY spPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTyrProI 1264  
Db ATTTTGCACACCCCAATATCATTTCTTGTGTTATGTAGTCAATTTCCACCTGGCCCA 3568  
QY leLysIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheV 1284  
Db TCAAAATTCAGATTGCCCTCTCTCAGCACAGCAATTTATGGACCTGCTGAGTCTTCG 3628  
QY alThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTyrLysIleL 1304  
Db TGACTTTCTTCAGTGTGTTCTCTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3688  
QY ysGlnSerCysTyrPheAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetA 1324  
Db AACAAAGTTGTTGGCTCCAGACGCTAGAGAGCAATCTCTTCGAGAGATGCAACAGATGG 3748  
QY laSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProAspL 1344  
Db CCAGCCCTGCTTTCCTCTGTAATGTCCCTTTGGAAACAGATGAGAGCCCTCCTGATC 3808  
QY euIleGlyGlySerIleLysThrValProLysProIleAlaLeuGluProCysPheGlyA 1364  
Db TTATTGGGGGAGTATTAAGACTGTTCCCAACCCCATTTGCATGGAGCCCTGTTTGGCA 3868  
QY snLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProp 1384  
Db ACAAGGCCGTGCTCTCTGTTGTTGAGGCTCCCTCGAGGCTCGGCTGGCATCCCTC 3928  
QY roProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetP 1404  
Db CTCTCTGGCAGTCAAGTCTGCTGCTGGCCAGGCCCCCTGGTGGACATTTCTCAGCAGATGC 3988  
QY roIleValTyrLysGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaG 1424  
Db CGATAGTGTACAGGAGAGTCAAGAGCCGTGAGAACCCGGAACCGAGCAGCCCCCTGCAC 4048  
QY lnProGlyThrCysIle 1429  
Db AGCTGGGACCTGCATC 4065  
RESULT 6  
ID AA291921 standard; cdna; 4072 BP.  
XX AA291921;  
AC AA291921;  
XX

DT 08-JUN-2000 (first entry)  
 XX Human mahogany protein coding sequence #3.  
 XX Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
 XX weight regulation; cell therapy; body weight disorder; cachexia;  
 XX anorexia; hyperpigmentation; increased metabolic rate disorder;  
 XX hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
 XX Homo sapiens.  
 XX WO200005373-A2.  
 XX 03-FEB-2000.  
 XX 21-JUL-1999; 99WO-US016484.  
 XX 21-JUL-1998; 98US-0093630P.  
 XX 20-OCT-1998; 98US-0104978P.  
 XX 03-FEB-1999; 99US-00245041.  
 XX (MILL-) MILLENIUM PHARM INC.  
 XX Moore K, Nagle DL;  
 XX WPI; 2000-195103/17.  
 XX P-PSDB; AAY81808.  
 XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
 XX treatment of body weight disorders.  
 XX Claim 1; Fig 19a; 188pp; English.  
 XX This sequence represents a human mahogany gene of the invention. The  
 XX mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 XX proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming  
 XX therapeutics; (iiii) as a source of diagnostic probes and primers for  
 XX detecting expression of mg genes or mutations, regulatory defects, in  
 XX this gene, or for isolation of related sequences; and (iv) in (cell-  
 XX based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
 XX identify other (extra)cellular products involved in weight regulation,  
 XX and to screen for agents that disrupt interaction between (ii) and other  
 XX macromolecules. The Ab are used to detect abnormal levels (or function)  
 XX of (ii) (for diagnosis, prognosis or monitoring of treatment); to  
 XX evaluate (ii)-expressing cells intended for cell therapy; and as  
 XX therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 XX mg polypeptide) are used to identify agents (A) that modulate mg  
 XX activity. (A) are potentially useful for the treatment of body weight  
 XX disorders, particularly obesity, cachexia or anorexia, or other  
 XX conditions associated with the mg gene such as hyperpigmentation,  
 XX hyperphagia and disorders that result in increased metabolic rate  
 XX SQ. Sequence 4072 BP; 1087 A; 940 C; 1045 G; 1000 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 4072  
 Score: 979.00 Matches: 1179  
 Percent Similarity: 99.66% Conservative: 0  
 Best Local Similarity: 99.66% Mismatches: 2  
 Query Match: 68.51% Indels: 4  
 DB: 3 Gaps: 0  
 US-09-787-097-12 (1-1429) x AA291921 (1-4072)  
 QY 87 AlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
 DB 256 GCGGGCGGCTGTCGGCTCGCGCGAGCCAGCAAGCAATGTGACCGCGGTGTGTC 315  
 QY 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126  
 DB 316 AACGGGGGTGCTGCAACCTGGACCGCCAGTGTGCTGCCCCCGCGGTGGGCGGC 375  
 QY 127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146

DB 376 GAGCAATGCCAGCACTCGGGGGCCGCTTCAGACTTAATGGATCTTCTGGGTTTGTGACA 435  
 QY 147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuLeuGluGlyGlnPro 166  
 DB 436 GATGGACCTGGAAATTAATAATCAAAACGAAGTCAGCGTGGCTCATTTGAAGACAGCCA 495  
 QY 167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186  
 DB 496 AATAGAATAATGAGACTTCGTTTCAATCATTTTGTCTACAGAGTGTAGTGGGACCAATTA 555  
 QY 187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206  
 DB 556 TATGTTTATGATGGGACTCAATTTATGACACCGCTAGTTGCTGCAATTTAGTGGCTCAT 615  
 QY 207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226  
 DB 616 GTTCCTCAGAGAGATGGCAATGAGACTGTCCCTGAGGTGTTCGCCATCAGGTTATGCC 675  
 QY 227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246  
 DB 676 TTGCTGCATTTTTTTAGTGAATGCTGCTTAATTTGACTGGATTAAATATTAATTACAGT 735  
 QY 247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266  
 DB 736 TTTGATATGTGTCCAAATAACTGCTCAGGCCGAGGAGAGTGAAGATCAGTAATAGCAGC 795  
 QY 267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286  
 DB 796 GA-TACTGTTGAAATGTGAATGTTCTGAAAACTGGAAGGTGAAGCATGTGACATTTCTCA 854  
 QY 286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgG1 306  
 DB 855 CTGTACAGACACACTGTGTTTCTCATCGAGGCATCTGCAATTTCAAGTGATGTCAGAGG 914  
 QY 306 YCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnG1 326  
 DB 915 ATGCTCTGCTTCTCAGACTGGCAGGTCCTGATGTTTCAAGTTCCTGTACACAGTAAACA 974  
 QY 326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346  
 DB 975 GTCAATTTGACTCGAGAGGAATATTTCACTTAAGTCTCCAGAGCATCTCATAAAGC 1034  
 QY 346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366  
 DB 1035 TGTGGTCAATGGAAACATTAATGTGGTGTGTGGAGGATATATGTTCAACCACTCAGATTA 1094  
 QY 366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386  
 DB 1095 TAAATGTTCTAGCGTATGACCTTGCTTCTAGGAGTGGCTTCCACTAAACCGTTCTGT 1154  
 QY 386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406  
 DB 1155 GAACAATGTGTTGTTAGATATGCTATTCTTTGGCATTATACAAGGATAAAATTTACAT 1214  
 QY 406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHis1 426  
 DB 1215 GTATGGAGGAAAAATTAATGATTC-AACTGGGAATGTGACCAATGATGTGAGAGTTTTTCA 1273  
 QY 426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446  
 DB 1274 TTCATATGAGTCATGGGTGTGTGACCCCTTAAGCCAAAGGAGCAGTATGTCAGTGGT 1333  
 QY 446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466  
 DB 1334 GGCACCTGCACACATTTTACACTGAAGAAATGGCGAGTGGTCATGCTGGTCACTCTTTG 1393  
 QY 466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486  
 DB 1394 GTCACTGCCCTCTCTATGGATATATAGCAATGTGCAGGAATATGATTTGGATAAGAAC 1453  
 QY 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerV 506

1454 Db CATGAGTATATTACACACCCAGGGTGGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513  
506 Qy altyrAspHisargThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526  
1514 Db TTATAGACCATAGGACCGAGGGCCCTATAGCTTATAGCTGCTGCTACAGGCTTTTCAGTGCCA 1573  
526 Qy snlystrArgLeuAlaAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546  
1574 Db ATAAAGTACCGGCTTGCCAGATGATCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633  
546 Qy leleuLysAspSerArgPheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566  
1634 Db TTCTTAAGACAGCCGATTTTTCGGTTACTTGCACACAGCTGTGATGTGAGTGGAAACA 1693  
566 Qy etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586  
1694 Db TGTGCTGTTGGGGGAAACACACACATGACACATCTATGAGCCATGGCGCCAAATGCT 1753  
586 Qy heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606  
1754 Db TCTCTTCAGATTCATGGCTATGACATTCGCTGTGACCGCTGGTCAGTGTCTCCAGAC 1813  
606 Qy roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626  
1814 Db CTGATCTCCACCATGATGTCACAGATTGGGCATTCAGCAGTCTTACACACAGACCA 1873  
626 Qy etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerg 646  
1874 Db TGTATGTTGCTGGTGTTCATAGTCTCTCCTCAGCGACATCTGGTATTCACCTGG 1933  
646 Qy luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
1934 Db AACAGTGTATGCGCATCGAGTGAAGCGCTGTGTTACGACGAGACCTGGTATTCGGT 1993  
666 Qy ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686  
1994 Db GTGTGTGGAAACACAGGTCGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAAG 2053  
686 Qy luGluLysLeuLysSerGluCysPheSerIlyArgThrLeuAspHisAspArgCysAspG 706  
2054 Db AAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113  
706 Qy lnhisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726  
2114 Db AGCACACAGATTGTTACAGCTGCACACCCAAACCAATGACTGCCACTGCTGCAATGACC 2173  
726 Qy isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746  
2174 Db ATTGTGTCCCGAAGAACACACAGCTGCTCAGAGGCCAGATCTCCATTTTATGATGAGA 2233  
746 Qy snCysProLysAspAsnProMetTyrCysAsnLysLysThrSerCysArgSerCysA 766  
2234 Db ATTGCCCCCAAGGATAACCCCTATGTTACTGTGTACAGAGAACACAGCTGCGAGAGCTGTG 2293  
766 Qy laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786  
2294 Db CCCTGGACCAAGAACTGCGAGTGGAGCCCGGAATCAGGAGTGCAATTGCCCTGCCCGAAA 2353  
786 Qy snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaI 806  
2354 Db ATATCTGTGCAATGGCTGGCCATTTGGTGGAAACTCATGTTTGAATAATTACTACTGCCA 2413  
806 Qy ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerI 826  
2414 Db AGGAAATTTATGACATGCTTAATGTTCTGTAGGAAACCAATGCGCTTTTGGCTTCTC 2473  
826 Qy euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerg 846  
2474 Db TTACACCCCAAGAAAGGTAGAATTTGCTCTTAAGCAGCTGCCAATAATGACATCATCTC 2533  
846 Qy lnsrMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866  
2534 Db AGAGCATGTCCAAAGCTCACCTTAACCCCATGGGTGCGCCCTTCGGAAGATCAATGTGCTCT 2593

866 Qy YrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886  
2594 Db ACTGGTCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATCCGCTCTG 2653  
886 Qy luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906  
2654 Db AGCCCACTGATGCTGGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGG 2713  
906 Qy laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926  
2714 Db CTGCAACTCTGCATCAACCCACTCAATGTTAGTGTCTGTGAAGGCTTCAACACACAGTG 2773  
926 Qy laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946  
2774 Db CTAAGCAGTCCGGACACCATGTCCTTTGAGGACAGCATGTGAGATTTGCACGACGCGCA 2833  
946 Qy erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966  
2834 Db GCTCTGAGTGCATGTGTGTGAGCAGACATGAAGCAGTGTGTGGACTCCAATGCCTATGTGG 2893  
966 Qy laSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986  
2894 Db CCTCCTTCCCTTTTGGCCAGTGTATGATGGTATACGATGAGCACCTGCCCCCTGAAA 2953  
986 Qy snCysSerGlyTyrCysThrCysSerHisCysLeuGlnProGlyCysGlyTyrCysT 1006  
2954 Db ATTGTTCAAGCTACTGTACCTGTAGTCAATGTTTGGAGCAACAGGCTGTGGTGTGTA 3013  
1006 Qy hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValL 1026  
3014 Db CTGATCCACAGCAATPACTGGCAAGGAAATGCTATAGAGGTTCTCTATAAAGGACCACTGA 3073  
1026 Qy ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerm 1046  
3074 Db AGATGCCCTTCGACAGCCCTTACAGGAAATTTCTATCCACAGCCCTCTCTCAATTCACGCA 3133  
1046 Qy etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066  
3134 Db TGTGTTCTAGAGGACAGACAGATACAACTGGTCTTTTCACTCTGCCAGCTTGCCAAATGCA 3193  
1066 Qy snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086  
3194 Db ACGGCCACAGTAAATGCAATCAATCAGACATCTCTGAGAAGTGTGAGAACCTGACCCACAG 3253  
1086 Qy llyHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106  
3254 Db GCAAGCACTGGAGACCTTGCATATCTGCTTCTACGGTGTATCCCAATGAGAGGAAAT 3313  
1106 Qy ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126  
3314 Db GTCAGCCATGCAAGTGGCAATGGGCACGCTCTCTGTGCAACACCAACACGCGGCAAGTCT 3373  
1126 Qy heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146  
3374 Db TCTGCACCAACAAAGGGCTCAAGGGGACGAGTCCAGCTATGTGAGTGAATAATCAT 3433  
1146 Qy YrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166  
3434 Db ACCAAGGAACCCCTCTCAGAGAAACATGTTATTATATCTCTCTTATTGACTATCAGTTCA 3493  
1166 Qy hrPheSerLeuSerGlnGluAspArgTyrTyrThrThrAlaIleAsnPheValAlaThrP 1186  
3494 Db CCTTTAGTCTATCCAGGAGATGATGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553  
1186 Qy roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206  
3554 Db CTGAGAAACAAACAGGGATTGGACATGTTTCATCAATGCTTCCAGAAATTTCAACCTCA 3613  
1206 Qy snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226  
3614 Db ACATCACCTGGCTGCCAGTTTCTCAGCTGGAAACCCAGGCTGGAGAAAGAGATGCTGTG 3673



QY 1226 alSerLysThrAsnIleGlyLeuValSerPheSerAsnGluLysPheAspPheA 1246  
 DB 3674 TTTCAAAACCAACATTAAGAGTACAAAGATAGTTTCTCTAATGAGAGTTGATTTC 3733  
 QY 1246 rGanHisProAsnIleThPhePheValTyValSerAsnPheThrProIleLysI 1266  
 DB 3734 GCAACACCAACCAATATACATTTCTTTTATGTCAGTAATTTTCACCTGGCCCATCAAAA 3793  
 QY 1266 leGln 1267  
 DB 3794 TTCAG 3798  
 RESULT 7  
 AA252274  
 ID AA252274 standard; cDNA; 3597 BP.  
 XX  
 AC AA252274;  
 DT 18-JUL-2000 (first entry)  
 XX  
 DE Human soluble attractin-1 cDNA.  
 XX  
 KW Human; soluble attractin-1; immune response; macrophage; monocyte;  
 T cell; immunostimulant; immunosuppressed patient; cancer;  
 immunodeficiency syndrome; transplant; autoimmune disease; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..3597  
 FT /\*tag= a  
 FT /product= "Soluble attractin-1"  
 XX  
 PN W0200015651-AL.  
 XX  
 XX 23-MAR-2000.  
 XX  
 XX 14-SEP-1999; 99WO-US020948.  
 XX  
 XX 14-SEP-1998; 98US-0100137P.  
 XX  
 PA (DAND ) DANA FARBER CANCER INST INC.  
 XX  
 PI Duke-Cohan JS, Schlossman SP;  
 XX  
 XX WPI; 2000-271373/23.  
 DR P-PSDB; AAY70689.  
 XX  
 XX Isolated nucleic acids encoding human attractin polypeptides useful for  
 enhancing immune responses.  
 XX  
 XX Claim 3; Fig 8; 120pp; English.  
 XX  
 CC The patent discloses four forms of human attractin polypeptides which  
 enhance immune response by promoting macrophage and monocyte spreading in  
 the presence of T cells. These include soluble attractin-1 and -2 and  
 membrane attractin-1 and -2. These various forms of attractin are encoded  
 by alternatively spliced mRNA molecule transcribed from a single gene.  
 CC The present sequence is a cDNA encoding soluble attractin-1 obtained from  
 activated T cell and human foetal liver libraries. This sequence can be  
 used to enhance immune response in immunosuppressed patients such as  
 those undergoing chemo- and radio-therapy treatment for cancer or those  
 suffering from common variable immunodeficiency syndrome. The proteins  
 may also be used to screen modulators (agonists and antagonists) of  
 immune responses which may also be used to regulate immune reactions.  
 CC Attractin antibodies can be used to inhibit immune response in transplant  
 recipients or patients afflicted with autoimmune disease  
 XX  
 SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 0 Length: 3597  
 Score: 874.00 Matches: 1160

Percent Similarity: 99.49% Conservative: 0  
 Best Local Similarity: 99.49% Mismatches: 3  
 Query Match: 61.16% Indels: 6  
 DB: 3 Gaps: 0  
 US-09-787-097-12 (1-1429) x AA252274 (1-3597)  
 QY 105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTyr 124  
 DB 91 TGTGTCAACGGCGGTCTGCAACCTGGCACCGCCAGTGGTCTGCCCGCGGCTGG 150  
 QY 125 ValGlyGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPhe 144  
 DB 151 GTGGCGAGCAATCCAGCACTGCGGGGGCGGCTTCAGACTAACTGGATCTTCTGGGTTT 210  
 QY 145 ValThrAspGlyProGlyAsnTyLysTyLysThrLysCysThrTyrLeuIleGluGly 164  
 DB 211 GTGACAGATGGACCTGGAAATTATAATACAAACGAAGTGCAGTGGCTCATTTGAAGGA 270  
 QY 165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTyrAsp 184  
 DB 271 CAGCCAAATAGATAATAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGAC 330  
 QY 185 HisLeuTyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaIlePheSerGly 204  
 DB 331 CATTTATATGTTATGATGGGACTCAATTTATGCACCGCTAGTTGCTGCATTTAGTGGC 390  
 QY 205 LeuIleValProGluArgAspGlyAsnGluThrValProGluValAlaIleThrSerGly 224  
 DB 391 CTCATTTCTCTGAGAGAGATGGCAATGAGACTGTCTCCTCAGGTTGTGCCACATCAGGT 450  
 QY 225 TyrAlaLeuLeuHisPhePheSerAspAlaIleTyrAsnLeuThrGlyPheAsnIleThr 244  
 DB 451 TATGCCCTGTGTCATTTTGTAGTATGCTGCTTATAATTTGACTGGATTATATATTACT 510  
 QY 245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsn 264  
 DB 511 TACAGTTTGTATGTCTCCAAATACTGCTCAGCGCGAGAGAGTGTAGATCAGTAAT 570  
 QY 265 SerSerGlu-ThrValGluCysGluCysSerGluAsnTyrPlyGlyGluAlaCysAspIle 284  
 DB 571 AGCAGCGA-TACTGTGTAATGTGAATGTTCTGAAACTGGAAAGGTGAACATGTGCAT 629  
 QY 284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304  
 DB 630 TCCCTCATGTACAGACCACTGTGTTTCTCTCATCGAGGCATCTGCATTTCAAGTATGT 689  
 QY 304 IArgGlyCysSerCysPheSerAspTyrGlnGlyProGlyCysSerValProValProAl 324  
 DB 690 CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGCTCTGGATGTTTCAGTTCCTGTACCAGC 749  
 QY 324 aAsnGlnSerPheThrArgGluTyrSerAsnLeuLysLeuProArgAlaSerHi 344  
 DB 750 TAACCATGCTATTTGGACTCGAGAGAAATTTAACTTAAAGCTCCCGAGACATCTCA 809  
 QY 344 sLysAlaValValAsnGlyAsnIleMetTyrValValGlyIleTyrMetPheAsnHis 364  
 DB 810 TAAAGCTGTGGTCAATCGAAACATTTATGTGGTGTGGAGGATATATGTTCAACCATC 869  
 QY 364 rAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTyrLeuProLeuAsnAr 384  
 DB 870 AGATTATACATGGTCTTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCCATAAACCG 929  
 QY 384 gSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysI 404  
 DB 930 TTCTGTGAACAATGTGGTGTGTAGATATGCTCATTTGGCATTTATACAGGATATAAT 989  
 QY 404 eTyrMetTyrGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424  
 DB 990 TTATCATGTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGATGTAGAGTTT 1048  
 QY 424 heHisIleHisAsnGluSerTyrValLeuLeuThrProLysAlaLysGluGlnTyrAlav 444



Db 1049 TTCACATTTCATATGAGTCATGGGTCTGTTGACCCCTTAAGCAAGAGGACGATGCGAG 1108  
Qy 444 alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464  
Db 1109 TGGTTGGGCACTCTGCACACATTGTTACACTGAAGAATGCCGAGTGGTCATGCTGGTCA 1168  
Qy 464 lePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGlnLysAspLeuAspL 484  
Db 1169 TCCTTGGTCACCTGCGCTCTCTATGGATATATAAGCAATGTGCAGGAATATGATTGGATA 1228  
Qy 484 yAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyrGlyHis 504  
Db 1229 AGAACACATGGAGATATTACACACCCAGGGTCCCTTGTGAAGGGGGTTACGGCCATA 1288  
Qy 504 eSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS 524  
Db 1289 GCGAGTGTTCAGCACCATAGGACAGGGGCCCTATACGTTTCATGGTGCTACAGAGCTTICA 1348  
Qy 524 eAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetT 544  
Db 1349 GTGCCAATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGT 1408  
Qy 544 rpThrIleLeuLysAspSerArgPheArgTyrIleHisThrAlaValIleValSerG 564  
Db 1409 GGACCAATCTTAAGGACAGCCGATTTTCCGTTACTTGCACACAGCTGTGATGAGTG 1468  
Qy 564 lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584  
Db 1469 GAACCATGTGGTGTGGGGAACACACACATGACACATCTATGAGCCATGGCGCA 1528  
Qy 584 yScysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuP 604  
Db 1529 AATGCTCTCTTCAGATTTCATGGCCATGACATTCGCTGTGACCGCTGTCAGTGCTTC 1588  
Qy 604 roArgProAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsn 623  
Db 1589 CCAGACTGATTC-CACCATGATGTCAACAGATTGGCCATTCCAGCAGTCTTACACAC 1647  
Qy 624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe 643  
Db 1648 AGCACCATGTATGTGTGGTGGTTTCAATAGTCTCTCTCTCAGCGACATCTCGGTATTC 1707  
Qy 644 ThrSerGlnGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaIleGlyProGly 663  
Db 1708 ACCTCGGAACAGTGTATCGGCATCGGAGTGAAGCGCTGTGTTAGCAGCAGACTGGT 1767  
Qy 664 IleArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683  
Db 1768 ATTCTGGTGTGTGGAAACACAGGCTGTCTCAGTGTATCTCTGGGGCGCTGGCAACTGAT 1827  
Qy 684 GluGlnGlnGlyLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703  
Db 1828 GAACAAGAGAAAGTTAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGA 1887  
Qy 704 CysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723  
Db 1888 TGTCAGCAGCACACAGATTGTTACAGCTGCACAGCCACCAACATGACTGCCACTGGTGC 1947  
Qy 724 AsnAspHisCysValProArgAsnHisSerCysSerGluGlnIleSerIlePheArg 743  
Db 1948 AATGACCAATGTGTCCCAAGAACACACAGCTGCTCAGAAGGCCAGATCTCCATTTTAGG 2007  
Qy 744 TyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysIleThrSerCysArg 763  
Db 2008 TATGAGAAATGTCCCAAGAGATTAACCTATGTACTCTGTAAACAAGAACCCAGCTGCAGG 2067  
Qy 764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783  
Db 2068 AGCTGTGCCCTGGACACAGAACTGCCAGTGGAGCCCGCAATCAGGAGTGCAATGCCCTG 2127  
Qy 784 ProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThr 803  
Db 2128 CCGGAAATATCTGTGGCATTTGGCTGGCATTTGGTTGGAAACTCATCTCTTTGAAATTA 2187

Qy 804 ThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823  
Db 2188 ACTCCCAAGGAGAATATGACAAATGCTAAATTTGTTGTAGGAACCAATATGCCCTTTTG 2247  
Qy 824 AlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGln 843  
Db 2248 GCTTCTCTTACACCCAGAGAAGGTAGAAATTTCTCTCTTAAGCAGCTGCGAATAATGCAG 2307  
Qy 844 SerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsn 863  
Db 2308 TCATCTCAGACATGTCCAAAGCTCACCTTAACCCATGGGTGGCCCTTCGGAAGATCAAT 2367  
Qy 864 ValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMet 883  
Db 2368 GTGTCTTACTGCTGCTGGGAAGATATGTCCCATTTTACAATAATAGTTTACTACAGTGATG 2427  
Qy 884 ProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGly 903  
Db 2428 CCGTCTGAGCCCACTGATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGA 2487  
Qy 904 LeuLysAlaLalaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsn 923  
Db 2488 CTGAAGGCTGCACCTGCATCACCCACTCAATGGTAGTGTCTGTGAAGCCCTGCAAC 2547  
Qy 924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943  
Db 2548 CACAGTGTGAAGCAGTGCCGACACCATGTGCTTGAGGACAGCATGTGGAGATTGCACC 2607  
Qy 944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963  
Db 2608 AGCGGCACTGTAGTGCATGTGGTGACGCAACATGAAGCAGTGTGTGGACTTCCAAATGCC 2667  
Qy 964 TyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983  
Db 2668 TATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAAATGATAGATAGACACCTGCCCC 2727  
Qy 984 ProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGly 1003  
Db 2728 CCTGAAATTTGTTCAGGCTACTGTACCTGTAGTCAATTTGTTGGAGCAACCCAGGCTGGC 2787  
Qy 1004 TrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGly 1023  
Db 2788 TGGTGTACTGATCCACCAATCTGGCAAGGGAATGCAATAGAGGGTTCTTATAAAGGA 2847  
Qy 1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043  
Db 2848 CCAGTGAAGATGCTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGCTCAAT 2907  
Qy 1044 SerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCys 1063  
Db 2908 TCCAGCATGTGTAGAGGACAGAGATACAACATGGTCTTTCTATCTACTGTCCAGCTTGC 2967  
Qy 1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083  
Db 2968 CAATGCAACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAAAGTGTGACAACTG 3027  
Qy 1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103  
Db 3028 ACCACAGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAACCAATGGA 3087  
Qy 1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123  
Db 3088 GGGAAATGTACCAATGCATGCAATGGGACCGCTCTCTGTGCAACACACACACGGGC 3147  
Qy 1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143  
Db 3148 AAGTCTTCTGCACCAACCAAGGGGCTCAAGGGGACGAGTGCCAGCTATGTGAGGTAGAA 3207  
Qy 1144 AsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyr 1163  
Db 3208 AATCGATACCAAGGAAACCCCTCTCAGAGGAACATGTTATTATATATCTCTCTTATTGACTAT 3267

CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3597 BP; 951 A; 822 C; 898 G; 926 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 0 Length: 3597  
 Score: 874.00 Matches: 1160  
 Percent Similarity: 99.49% Conservative: 0  
 Best Local Similarity: 99.49% Mismatches: 3  
 Query Match: 61.16% Indels: 6  
 DB: Gaps: 5

US-09-787-097-12 (1-1429) x AAS72657 (1-3597)

QY 105 CysValAsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrp 124  
 Db 91 TGTCTCAACGGCGTCTGCTCAACCTTGGCCACCGCCAGTGGCTCTGCCCGCGCTGG 150  
 QY 125 ValGlyGlnGlnCysGlnHisCysGlyGlyValArgPheArgLeuThrGlySerSerGlyPhe 144  
 Db 151 GTGGCGAGCAATGCCAGCACTCTGGGGGGCGCTTACAGCTAACTGAGTCTTCTGGTTT 210  
 QY 145 ValThrAspGlyProGlyAsnTyrlsTyrlsTyrlsCysThrTrpLeuLeuGluGly 164  
 Db 211 GTGACAGATGGACCTGGAAATTTATAATACAAACGAAGTGCAGCTGCTCATTTGAAGA 270  
 QY 165 GlnProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAsp 184  
 Db 271 CAGCCAAATAGAAATATGAGACTTCGTTCAATCATTTTCTACAGAGTGTAGTTGGGAC 330  
 QY 185 HisLeuTyrlsValTyrlsAspGlyAspSerIleTyrlsAlaProLeuValAlaAlaPheSerGly 204  
 Db 331 CATTTATATGTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGATTTAGTGGC 390  
 QY 205 LeuLeuValProGluArgAspGlyAsnGluThrValProGluValAlaAlaThrSerGly 224  
 Db 391 CTCATTGTTCTCTGAGAGAGATGGCAATGAGACTGTCCCTGAGGTGTGTCACATCAGGT 450  
 QY 225 TyrAlaLeuLeuHisPhePheSerAspAlaAlaTyrlsAsnLeuThrGlyPheAsnIleThr 244  
 Db 451 TATGCTTGTGTCATTTTATGATGCTGCTTATTAATTTGACTGGATTTATATTAAT 510  
 QY 245 TyrSerPheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysIleSerAsn 264  
 Db 511 TACAGTTTGTATATGTTCTCAAAATACTGCTGAGCGGAGAGAGTGTAAAGATCAGTAAT 570  
 QY 265 SerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIle 284  
 Db 571 AGCAGCGA-TACTGTTGAATGTGAATGTTCTGAATACTGGAAGGTGAACATGTGCAT 629  
 QY 284 eProHisCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspVa 304  
 Db 630 TCCTCACTGTACAGACAACTGTGGTTTCTCTCATCGAGGCATCTGCAATTCAGATGATGT 689  
 QY 304 lArgGlyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProAla 324  
 Db 690 CAGAGGATGCTCTCTGCTCTCAGACTGGCAGGGTCTCTGGATGTTTCTGTTCTGTACAGC 749  
 QY 324 aAsnGlnSerPheTrpThrArgGluTyrlsSerAsnLeuLeuProAlaSerHis 344  
 Db 750 TAACCACTCAATTTGGACTCGAGAGGAATATTTCACTTAAGGCTCCCGAGAGCATCTCA 809  
 QY 344 sLysAlaValValAsnGlyAsnIleMetTrpValValGlyTyrlsMetPheAsnHisSe 364

QY 1164 GlnPheThrPheSerLeuSerGlnGluAspArgTyrlsThrAlaIleAsnPheVal 1183  
 Db 3268 CAGTTCACCTTTAGTCTATCCAGGAAGATGCTGCTATTACACAGCTATCAATTTTGTG 3327  
 QY 1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203  
 Db 3328 GCTACTCTCTGACGACAAACAGGGATTTGGACATGTTCTATCATGTCCTCCAGAAATTC 3387  
 QY 1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMet 1223  
 Db 3388 AACCTCAACATCACCCTGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3447  
 QY 1224 ProValValSerLysThrAsnIleLysGluTyrlsAspSerPheSerAsnGluLysPhe 1243  
 Db 3448 CTTGTTGTTTCAAAACCAACATTAAGGAGTACAAAGATAGTTTCTTAATGAGAAGTTT 3507  
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrlsValSerAsnPheThrTrpPro 1263  
 Db 3508 GATTTTCGCAACCAACCAATATCACTTCTTTGTTTATGTCAGTAATTTCACTGGGCC 3567

QY 1264 IleLysIleGln 1267  
 Db 3568 ATCAAAATTCAG 3579

RESULT 8  
 AAS72657  
 ID AAS72657 standard; cDNA; 3597 BP.  
 XX  
 AC AAS72657;  
 XX  
 DT 13-FEB-2002 (first entry)  
 XX  
 DE DNA encoding novel human diagnostic protein #8461.  
 XX  
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic; ss.  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN W0200175067-A2.  
 XX  
 PD 11-OCT-2001.  
 XX  
 XX 30-MAR-2001; 2001WO-US008631.  
 XX  
 XX 31-MAR-2000; 2000US-00540217.  
 PR  
 PR 23-AUG-2000; 2000US-00649167.  
 XX  
 XX (HYSE-) HYSEQ INC.  
 XX  
 XX Drmanac RT, Liu C, Tang YT;  
 PI  
 XX WPI: 2001-639362/73.  
 DR  
 DR P-PSDB; ABG08470.  
 XX  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX  
 PS Claim 1; SEQ ID NO 8461; 103pp; English.  
 CC  
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders

810 TAAAGCTGGTCAATGGAAACATTTATGTGGGTTGTTGGAGGATATATGTTCAACACCTC 869  
364 rAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAr 384  
870 AGATTATAACATGGTTCTAGCGTATGACCTTGCTTCTAGCGAGTGGCTTCCACTAAACCG 929  
384 gSerValAsnAsnValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysI 404  
930 TCTGTGAACAATGTGGTTGTAGATATGTCATCTTTGGCCATTATACAAGGATAAAAT 989  
404 eTyrMetTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValP 424  
990 TTAACATGTATGGAGGAAAATTAATGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTT 1048  
424 heHisIleHisAsnGluSerTrpValLeuLeuThrTrpProLysAlaLysGluGlnTyrAlaV 444  
1049 TTAACATTAATGATGATGGGTGTGTTGACCCCTTAAGCCAAAGGAGCAGTAGTCAG 1108  
444 aValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI 464  
1109 TGGTTGGGCACTTCGCACACATTTGTACACTGAGNAATGGCCGAGTGGTCATGCTGTCA 1168  
464 lPheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspL 484  
1169 TCTTTGGTCACTGCCCTCTCTATGGATATATAGCAATGTGCAGGAATATGATTTGGATA 1228  
484 yAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisS 504  
1229 AGAACACATGGAGTATATTACACCCAGGGTGCCCTGTGCAAGGGGTACGGCCATA 1288  
504 eSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS 524  
1289 GCAGTGTAAACACCATAGGACAGGGCCCTATAGCTTCATGGTGGCTACAGGCTTTCA 1348  
524 eAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetT 544  
1349 GTGCCAATAAGTACCGGCTTGAGATGATCTTACCGATATATGTTGGATACCCAGATGT 1408  
544 rThrIleLeuLysAspSerArgPheArgTyrIleuHisThrAlaValIleValSerG 564  
1409 GGACCACTTTAAGGACACCGCATTTTCCGTTACTTCACACAGCTGTGATAGTAGTG 1468  
564 lYThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL 584  
1469 GAACCATGCTGGTGTGTTGGGGAAACACACACAATGACACATCTATGAGCCATGCGCCA 1528  
584 yCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuP 604  
1529 AATGCTTCTCTCAGATTTTCATGGCCTATGACATTCGCTGACCGCTGGTCAGTGCCTC 1588  
604 rArgProAspLeuHisHisAspValAsnArgPheGlyHisSer-AlaValLeuHisAsn 623  
1589 CCAGACCTGATTC-CACCATGATGTCAACAGATTTGGCCATTTCCAGCAGTCTTACACAAC 1647  
624 SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleuValPhe 643  
1648 AGCACCATGATGTGTTCCGTTGGTTTCAATAGTCTCCTCCTCAGGACATCTCGTATTC 1707  
644 ThrSerGluGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaGlyProGly 663  
1708 ACCTCGNAACAGTGTGATCGCATCGAGTGAAGCGCTGTTTATGACGACGACCTGGT 1767  
664 lLeArgCysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAsp 683  
1768 ATTCCGTGTGTGGAAACACAGGCTGCTCTCAGTGTATCTCGTGGCGCTGGCAACTGAT 1827  
684 GluGlnGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArg 703  
1828 GAACAAGAGAAAAGTTAAATCAGATGTTTTTCCAAAAGAACTCTTTGACCATGACAGA 1887  
704 CysAspGlnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCys 723  
1888 TGTGACCAACACACATTTTACAGCTGCACAGGCCAACACCAATGACTGCCACTGGTGC 1947

724 AsnAspHisCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArg 743  
1948 AATGACCATTTGTGCCCGAGAAACACAGCTGCTCAGAGGCCAGATCTCCATTTTAGG 2007  
744 TyrGluAsnCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArg 763  
2008 TATGAGAAATTCGCCCAAGGATTAACCTTATGTACTTGTAAACAAGAAGACAGCTGCAGG 2067  
764 SerCysAlaLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeu 783  
2068 AGCTGTGCCCTGGACCAAGAACTGCAGTGGAGGCCCGAATCAGGAGTGCATTTGCCCTG 2127  
784 ProGluAsnIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThr 803  
2128 CCGAAAATATCTGTGGCATTTGGCTGGCATTTGTTGGAAACTCATGTTTGAATAATTACT 2187  
804 ThrAlaLysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeu 823  
2188 ACTGCCAAGGAGAAATTAAGCAATGCTTAATGTTCTGTAGGAACCAACAATGCCCTTTG 2247  
824 AlaSerLeuThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGln 843  
2248 GCTTCTCTTAAACCCCAAGAAAGGTAGAATTTGCTCCTTAAGCAGCTGCGAATAATGCAG 2307  
844 SerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsn 863  
2308 TCATCTGAGGATGTCACAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAAT 2367  
864 ValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMet 883  
2368 GTGTCTTACTGCTGGGAAGATATGTCCTCAATTTACAAATAGTTTACTACAGTGGATG 2427  
884 ProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGly 903  
2428 CCGTCTGAGCCCGAGTATGCTGGATTTCTGTGGAAATTTATCAGAACCCAGTACTCGGGA 2487  
904 LeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsn 923  
2488 CTGAAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCCCTGCAAC 2547  
924 HisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThr 943  
2548 CACAGTCTTAAGCAGTCCCGACACCATGTGCTTGGAGCAGCATGTGGAGATGGACCC 2607  
944 SerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963  
2608 AGCGGCAAGCTCTGAGTGCATGTGGTGACAGCAACATGAGCAGTGTGTGGACTCCAATGCC 2667  
964 TyrValAlaSerPhePropheGlyGlnCysMetGluTrpTyrThrMetSerThrCysPro 983  
2668 TATGTGGCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCCTGCCCC 2727  
984 ProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGly 1003  
2728 CCGAAAATTTGTTGAGGCTACTGTACTGTAGTCAATTTGCTGGAGCAACAGGCTGTGGC 2787  
1004 TrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGly 1023  
2788 TGGTGTACTGATCCAGCAATATCTGGCAAGGGAATGTCATAGAGGCTTCTCTATAAAGGA 2847  
1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043  
2848 CCAGTGAAGATGGCTTCGCAAGGCCCTTACAGGAATTTTCTATCCAGAGCCCTGCTCAAT 2907  
1044 SerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCys 1063  
2908 TCCAGCATGTGTCTAGAGGACAGCAGATACAATGGTCTTTTCTTCACTCTCCAGCTTGC 2967  
1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083  
2968 CAATGCAACGCCACACATAAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAGAAGCCTG 3027



Qy	265	SerSerGlu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAsp	11	284
Db	571	AGCAGCGCA-TACTGTTGAATGTGAATGTTCTGAAAACTCGAAAGGTGAACATGTGACAT	629	
Qy	284	eProHisCysThrAspAsnCySGlyPheProHisArgGlyVleCysAsnSerSerAsp	304	
Db	630	TCCTCACTGTACAGACACACTGTGGTTTCTCTCATCGAGGCATCTGCAATTCAGTGTATGT	689	
Qy	304	larGglyCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAl	324	
Db	690	CAGAGGATGCTCCTGCTTCTCAGACTGGCAGGTCCTGGATGTTTCAAGTCTCTCTGTACAGC	749	
Qy	324	aAsnGlnSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHi	344	
Db	750	TAAACCACTCAATTTTGGACTCGAGAGGAATATCTTAACCTTAAAGCTCCCGAGACATCTCA	809	
Qy	344	sLysAlaValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHis	364	
Db	810	TAAAGCTGTGGTCAATGGAACCAATATGTGGGTGTGGAGGATATATGTTCAACCACTC	869	
Qy	364	rAspTyrAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnAr	384	
Db	870	AGATTAAACATGGTCTCAGCGTATGACCTTGCTTCTTAGGAGGTGGCTTCCACTAAACCG	929	
Qy	384	gSerValAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLys	404	
Db	930	TTCTGTGAACAATGTGGTGTGTAGATATGTCTTCTTGGCATTTATACAAGGATAAAAT	989	
Qy	404	eTyrMetTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuLeuArgValP	424	
Db	990	TTACATGTAATGGAGGAAAAATTGATTC-AACTGGAAATGAGCAATGAGTTCAGAGT	1048	
Qy	424	heHisIleHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaV	444	
Db	1049	TTCACATTCATATGAGTCATGGGTGTGTGGACCTTAAGCCAAAGGAGCATGTGCAG	1108	
Qy	444	alValGlyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValI	464	
Db	1109	TGGTTGGGCACCTCGCACACATTTTACACTGAAGAATGGCCGAGTGGTCATGCTGGTCA	1168	
Qy	464	lePheGlyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspL	484	
Db	1169	TCITTTGGTCACTGGCCCTCTCTATGGATATATAGCAATGTGCAGGAATATGATTTGGATA	1228	
Qy	484	ysAsnThrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHis	504	
Db	1229	AGAACACATGGAGTATATTACACACCCAGGCTGCCCTTGTGCAAGGGGTTACGGCCATA	1288	
Qy	504	erSerValTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheS	524	
Db	1289	GCACTGTTTACGACCATAGGACACAGGCCCTATACGTTCACTGTGGTGGCTCAAGGCTTCA	1348	
Qy	524	erAlaAsnLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMet	544	
Db	1349	GTGGCAATAGTACCGGCTTGACATGATCTCTACCCATATGATGTGATACCCAGATGT	1408	
Qy	544	rpThrIleLeuLysAspSerArgPheArgTyrLeuHisThrAlaValIleValSerG	564	
Db	1409	GGACCATTCCTTAAGCAGACCGCGATTTTCGTACTTCTGCAACAGCTGTGATAGTGA	1468	
Qy	564	lyThrMetLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaL	584	
Db	1469	GAACCATGCTGTGTTTGGGGGAAACACACACATGACACATCTATGAGCCATGGCCCA	1528	
Qy	584	ysCysPheSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuP	604	
Db	1529	AATGCTTCTCTCCAGATTTCATGCCCTATGACATGTGCTGTGACCGGTGTGATGTGCTTC	1588	
Qy	604	roArgProAspLeuHisIleAspValAsnArgPheGlyHisSer-AlaValLeuHisAsn	623	
Db	1589	CCAGCCTGATTC-CACCATGATGTCAACAGATTGGCCATTTCAGCAGTCTTACACAC	1647	
Qy	624	SerThrMetTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPhe	643	

Db 2728 CCTGAAATTTCTAGGCTACTGTACTGTAGTCAATTTGCTTGAGCAACGAGGCTGTGCG 2787  
 QY 1004 TTPCysThrAspProSerAsnThrGlyLysGlyCysIleGluGlySerTyrLysGly 1023  
 Db 2788 TGGTGTACTGATCCAGCATATCTGGCAAGGGAATGCATAGAGGTTCTCTATAAGA 2847  
 QY 1024 ProValLysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsn 1043  
 Db 2848 CCAGTGAAGATGCTTCCGCAAGCCCTTACAGGAATTTCTATCCAGAGCCCTGCTCAAT 2907  
 QY 1044 SerSerMetCysLeuGluAspSerArgTyrAsnTyrPheIleHisCysProAlaCys 1063  
 Db 2908 TCAGCATGTCTTAGAGGACAGACATACATCTGTCTTCTCATCTCAGTCTCAGTTGC 2967  
 QY 1064 GlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeu 1083  
 Db 2968 CAATGCAACGCCACATGAATGCAATCATCATCATCATCATCATCATCATCATCATCAT 3027  
 QY 1084 ThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGly 1103  
 Db 3028 ACCACAGGCAAGCACTGCGAGACTGATATCTGGCTTCTACGTCATCCCAATGGA 3087  
 QY 1104 GlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGly 1123  
 Db 3088 GGGAAATGTCAGCATGCAATGGCAGCGCTCTGTGCAACCAACACGAGGC 3147  
 QY 1124 LysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGlu 1143  
 Db 3148 AAGTGTCTTCTGCAACCAAGGCGCTCAAGGGGAGCGAGTGCAGCTATGTGAGGTAGAA 3207  
 QY 1144 AsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuLeuAspTyr 1163  
 Db 3208 AATGATACCAAGAAACCTCTCAGAGGACATGTTATATCTTCTTATGACTAT 3267  
 QY 1164 GlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheVal 1183  
 Db 3268 CAGTTACCTTTAGTCTATCCAGGAAGATGATCGCTATTACAGCTATCATCAATTTGTG 3327  
 QY 1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203  
 Db 3328 GCTACTCTCAGCAACAAACAGGATTTGGACATGTTTATCATATCCCTCCAGAAATTC 3387  
 QY 1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluMet 1223  
 Db 3388 AACTCAACATCCTCGGCTGCCAGTTCTCAGCTGGAACCCAGGCTGGAGAGAGATG 3447  
 QY 1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243  
 Db 3448 CCTGTTGTTTCAAAAACCAACATTAAGGAGTACAAAGATAGTTCTTCTAATCAGAAATTT 3507  
 QY 1244 AspPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrpPro 1263  
 Db 3508 GATTTTCGCAACCAACCAATATCACTTCTTGTGTTATGTGATTAATTCACCTGGCC 3567  
 QY 1264 IleLysIleGln 1267  
 Db 3568 ATCAAAATTCAG 3579  
 RESULT 10  
 ID AA291922 standard; cDNA; 2625 BP.  
 XX  
 AC AA291922;  
 XX  
 DT 08-JUN-2000 (first entry)  
 XX  
 DE Human mahogany protein coding sequence #4.  
 XX  
 KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
 XX

OS Homo sapiens.  
 XX WO200005373-A2.  
 PN 03-FEB-2000.  
 PD 21-JUL-1999; 99WO-US016484.  
 PF 21-JUL-1998; 98US-0093630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 99US-00245041.  
 XX (MILL-) MILLENIUM PHARM INC.  
 XX Moore K, Nagle DL;  
 PI WPI; 2000-195103/17.  
 DR P-PSDB; AAY81809.  
 XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.  
 XX Claim 1; Fig 20a; 188pp; English.  
 CC This sequence represents a human mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (ii) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (ii) for diagnosis, prognosis or monitoring of treatment; to  
 CC evaluate (ii)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate.  
 XX Sequence 2625 BP; 660 A; 605 C; 694 G; 666 T; 0 U; 0 Other;  
 SQ  
 Alignment Scores:  
 Pred. No.: 0 Length: 2625  
 Score: 496.00 Matches: 696  
 Percent Similarity: 99.43% Conservative: 0  
 Best Local Similarity: 99.43% Mismatches: 2  
 Query Match: 34.71% Indels: 4  
 DB: 3 Gaps: 0  
 US-09-787-097-12 (1-1429) x AA291922 (1-2625)  
 QY 87 AlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
 Db 256 GCGGCGGGTGTCTCGGCTCGCGCAGCCGAGGCAAGGAATGTGACCGCGCTGTGTC 315  
 QY 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTTPValGly 126  
 Db 316 AACGGCGGTGCTGCAACCTTGGCACCGGCAGTGGCTGTGCCCCCGGCTGGGTGGC 375  
 QY 127 GluGlnCysGlnHisCysGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146  
 Db 376 GAGCAATGCCACACTCGCGGGCGCTTCAGACTTAACGATCTTCTGCGTTGTGACA 435  
 QY 147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuIleGluGlyGlnPro 166  
 Db 436 GATGGACTGGAATTAATAATCAAAACGAATGACGTGGCTCATTTGAAGACAGCCCA 495  
 QY 167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186



```
Db 496 AATAGAAATATGAGACTTCGTTTCAATCATTTTGGCTACAGACTGTAGTTGGGACCATTTA 555
Qy TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206
Db 556 TATGTTTATGATGGGACTCAATTTATGACCGCTAGTTGCTGCAATTTAGTGGCCCTCAT 615
Qy ValProGluArgAspGlyAsnGluThrValProGluValAlaAlaThrSerGlyTyrAla 226
Db 616 GTTCCTGAGAGAGATGGCAATGAGACTGTCCTCGAGGTTGTTGGCCACATCAGGTTATGCC 675
Qy LeuLeuHisPhePheSerAspAlaLaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246
Db 676 TTGCTGCATTTTATGATGATGCTGCTTATATTTGACTGCAATTTAATTTACTTACAGT 735
Qy PheAspMetCysProAsnAsnCysSerGlyAaGGlyGluCysIleSerAsnSerSer 266
Db 736 TTTGATATGTGTCCTCAATTAACCTGCTCAGCCGAGGAGAGTGAAGATCAGTAATAGCAGC 795
Qy Glu-ThrValGluCysGluCysSerCysLeuThrPheGlyGluAlaCysAspIleProHi 286
Db 796 GA-TACTGTGTGAATGTGAATGTTCTGAATACTGGAAAGGTGAAGCATGTGACATTCCTCA 854
Qy sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgI 306
Db 855 CTGTACAGACAACTGTGGTTTCTCATCGAGGCATCTGCAATTCAGTGTGTCAGAGG 914
Qy yCysSerCysPheSerAspThrGlnGlyProGlyCysSerValProValProAlaAsnG 326
Db 915 ATGCTCTCTCTCTCAGACTGCGAGGCTCTGAGTGTTCAGTTCCTGACCTTACCAGCTTA 974
Qy nSerPheThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346
Db 975 GTCAATTTGGACTCGAGAGGATATCTTAACCTTAAGCTCCCGAGGATCTCATTAAGC 1034
Qy aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366
Db 1035 TGTGCTCAATGGAACCATTTATGTGGGTTGTTGGAGGATATATGTTCAACCACTCAGATTA 1094
Qy rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386
Db 1095 TAACTGGTTCTAGCGTATGACCTTCTGCTTCTAGGAGGTGGCTTCCACTAAACCGTCTGT 1154
Qy lAsnAsnValValAlaTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406
Db 1155 GAACCAATGTGGTTGTAGATATGTTCAATTTCTTTGGCATTTATACAGGATAAATTTACAT 1214
Qy tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426
Db 1215 GTATGGAGGAAAAATGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACA 1273
Qy leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446
Db 1274 TTCAATAGTGCATGGGTGTTGTTGACCCCTAAGCAAGGAGCAGTATGCACTGGTTG 1333
Qy lHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466
Db 1334 GGCACCTCTGCACACATTTGTACACTGAAGATGGCCGAGTGTCTATGCTGCTCATCTTTG 1393
Qy lHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsn 486
Db 1394 GTCACGTCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAGAACA 1453
Qy hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSerSerV 506
Db 1454 CATGGAGTATATTACACACCCAGGGTGCCCTTGTGCAAGGGGGTTACGGCCATACAGTG 1513
Qy alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526
Db 1514 TTATAGACCATAGGACCGGCGCTTATACGTTTCATGGTGGCTTACAGGCTTTCAGTGCCA 1573
Qy snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546
Db 1574 ATAGTACCGGCTTCGACGATGATCTCTACCGATATGATGTGATACCGAGTGTGGACA 1633
```

```
Qy 546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566
Db 1634 TTCTTAAGGACAGCCGATTTTCGTTACTTGCACACAGCTGTGATAGTGGTGAACCA 1693
Qy etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586
Db 1694 TGCTGGTGTGTTGGGGGAAACACACACAATGACATCTATGAGCCATGGGCCCAATGCT 1753
Qy heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTyrTrpSerValLeuProArgP 606
Db 1754 TCCTCTCAGATTTTCATGGCTTATGACATTCGCTGTGACCCGCTGGTCACTGTTCCAGAC 1813
Qy roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626
Db 1814 CTGATCTCCACCATGATGTCAACAGATTTGGCCATTTCAGCAGTCTTTACACACAGCACCA 1873
Qy etTyrValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646
Db 1874 TGTATGTGTCGGTGGTTCATATAGTCTCTCTCTCAGCAGATCTCTGTTATTCACCTCGG 1933
Qy luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666
Db 1934 AACAGTGTGATGCGATCGGAGTGAAGCCGCTTTTTCAGCAGCAGGACCTGGTATTCGGT 1993
Qy ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686
Db 1994 GTGTGTGGAACACACAGGTCGTCTCAGTGTATCTCGTGGGCGCTGGCAACTGATGAACAAG 2053
Qy luGluLysLeuLysSerGluCysPheSerLysArgThrLeuAspHisAspArgCysAspG 706
Db 2054 AAGAAAGTTTAAATCAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTGACC 2113
Qy lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726
Db 2114 AGCACACAGATTTGTTACAGCTGCACAGCAACCAATGATGCTGCTGCTGCTGCTGCTG 2173
Qy isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db 2174 ATTGTGTCGCGAGAACACACAGCTGCTCAGAGCCGAGATCTCAATTTTATGATGAGA 2233
Qy snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db 2234 ATTGCCCCAAGGATAACCTTATGCTACTGTAAAGAAAGACAGCTGCGAGGAGCTGTG 2293
Qy lateuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784
Db 2294 CCTGAGACCAAGAACTGCCAGTGGAGGCCCGGAATCAGGAGTGCATTTGCTGCTGCC 2349
RESULT 11
AAZ91919
ID AAZ91919 standard; cDNA; 6373 BP.
XX
AC AAZ91919;
XX
DT 08-JUN-2000 (first entry)
XX
DE Human mahogany protein coding sequence #1.
XX
KW Mahogany gene; human; mg gene; regulatory defect; gene therapy; obesity;
weight regulation; cell therapy; body weight disorder; cachexia;
anorexia; hyperpigmentation; increased metabolic rate disorder;
hyperphagia; Antiobesity; antianorexic; anticachexic; ss.
XX
OS Homo sapiens.
XX
Key Location/Qualifiers
FT misc_difference 448 /*tag= a
FT /*note= "represented in specification as : "
FT misc_difference 615 /*tag= b
FT /*note= "represented in specification as : "
```





Db 1937 TCTTGTGTGGCCAGCCCTGGTGGACATTTCTAGCGATGCCGATAGTGTACAGGA 1996  
QY 1409 ulysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIi 1429  
Db 1997 GAAGTCAGAGCGGTGAGAAACCGGAAGCAGCAGCCCCCTGCACAGCCTGGGACCTGCAT 2056  
QY 1429 e 1429  
Db 2057 C 2057  
RESULT 12  
ID AAS72660 standard; cDNA; 3490 BP.  
XX AAS72660;  
AC AAS72660;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #8464.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
XX WO200175067-A2.  
XX  
XX 11-OCT-2001.  
XX  
XX 30-MAR-2001; 2001WO-US008631.  
XX  
XX 31-MAR-2000; 2000US-00540217.  
XX  
XX 23-AUG-2000; 2000US-00649167.  
XX  
XX (HYSE-) HYSEQ INC.  
XX  
XX  
XX Drmanac RT, Liu C, Tang YT;  
XX  
XX WPI; 2001-639362/73.  
XX  
XX P-PSDB; ABG08473.  
XX  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in-  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits and to assess  
XX biodiversity.  
XX  
XX Claim 1; SEQ ID NO 8464; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
XX sequences. (I) is useful as hybridisation probes, polymerase chain  
XX reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
XX and in recombinant production of (II). The polynucleotides are also used  
XX in diagnostics as expressed sequence tags for identifying expressed  
XX genes. (I) is useful in gene therapy techniques to restore normal  
XX activity of (II) or to treat disease states involving (II). (II) is  
XX useful for generating antibodies against it, detecting or quantitating a  
XX polypeptide in tissue, as molecular weight markers and as a food  
XX supplement. (II) and its binding partners are useful in medical imaging  
XX of sites expressing (II). (I) and (II) are useful for treating disorders  
XX involving aberrant protein expression or biological activity. The  
XX polypeptide and polynucleotide sequences have applications in  
XX diagnostics, forensics, gene mapping, identification of mutations  
XX responsible for genetic disorders or other traits to assess biodiversity  
XX and to produce other types of data and products dependent on DNA and  
XX amino acid sequences. AA564197-AA594564 represent novel human diagnostic  
XX coding sequences of the invention. Note: The sequence data for this  
XX patent did not appear in the printed specification, but was obtained in  
XX electronic format directly from WIPO at  
XX ftp.wipo.int/pub/published\_pct\_sequences  
XX  
XX Sequence 3490 BP; 931 A; 772 C; 858 G; 929 T; 0 U; 0 Other;

Alignment Scores: 3.2e-245 Length: 3490  
Pred. No.: 289.00 Matches: 528  
Score: 98.69% Conservative: 0  
Percent Similarity: 98.69% Mismatches: 2  
Best Local Similarity: 20.22% Indels: 7  
Query Match: 5 Gaps: 0  
DB: 5  
US-09-787-097-12 (1-1429) x AAS72660 (1-3490)  
QY 137 ArgLeuThrGlySerSerGlyPheValThrAspGlyProGlyAsnTyrLysTyrLysThr 156  
Db 22 AGACTAACTGATCTTCTGGGTTTGTGACAGATGGACCTGGAATTAATAATACAAACG 81  
QY 157 LysCysThrTrpLeuIleGluGlnProAsnArgIleMetArgLeuArgPheAsnHis 176  
Db 82 AAGTGCACGTGGCTCATTTGAAGGCACGCCAATAGAATAATAGACTTCGTTTCAATCAT 141  
QY 177 PheAlaThrGluCysSerTrpAspHisLeuTyrValTyrAspGlyAspSerIleTyrAla 196  
Db 142 TTTGCTACAGAGTGTAGTGGGACCAATTTATATGTTATGATGGGACTCAATTTATGCA 201  
QY 197 ProLeuValAlaAlaPheSerGlyLeuIleValProGluArgAspGlyAsnGluThrVal 216  
Db 202 CCGCTAGTGTGCTGCAATTTAGTGGCCTCATTTGCTCAGAGAGATGGCAATGAGCTGTC 261  
QY 217 ProGluValValAlaThrSerGlyTyrAlaLeuLeuHisPheSerAspAlaAlaTyr 236  
Db 262 CCTGAGGTGTGGCCACATCAGGTTATGCTTGTGCAATTTTATGATGCTGCTTAT 321  
QY 237 AsnLeuThrGlyPheAsnIleThrTyrSerPheAspMetCysProAsnAsnCysSerGly 256  
Db 322 AATTTGACTGATTAATATTTACTTACAGTTTGTATGATGTGCCAAATAAATCTGCTAGGC 381  
QY 257 ArgGlyGluCysLysIleSerAsnSerSerGlu-ThrValGluCysGluCysSerGluAs 276  
Db 382 CGAGAGAGTGTAGATCAGTAATAGAGCGCA-TACTGTTGAATGTGAATGTTCTGAAA 440  
QY 276 nTrpLysGlyGluAlaCysAspIleProHisCysThrAspAsnCysGlyPheProHisAr 296  
Db 441 CTGGAAGGTGAAGCATGTGACATTCCTCACTACAGACAACTGTGTTTCTCTCATCG 500  
QY 296 gGlyIleCysAsnSerSerAspValArgGlyCysSerCysPheSerAspTrpGlnGlyPr 316  
Db 501 AGGCATCTGCAATTCAGTGTATGTCAGAGGATGCTCTCTGCTTCTCAGACTGGCAGGGTCC 560  
QY 316 oGlyCysSerValProValProAlaAsnGlnSerPheTrp-ThrArgGluGluTyrSerA 336  
Db 561 TGGATGTTTCACTCTGTACCAGCTAACCACTCATTTTGGGACTCGAGAGGAATATTCTA 620  
QY 336 sn-LeuLysLeu-ProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTrpVa 355  
Db 621 ACTTTAAAGCTCCGCCAGAGCATCTCATAAAGCTGTGTGTCATGGAACATTTATGTGGGT 680  
QY 355 lValGlyGlyTyrMetPheAsnHisSerAspTyrAsnMetValLeuAlaTyrAspLeuAl 375  
Db 681 TGTGGAGGATATATGTTCAACCACTCAGATTATACATGGTTCTAGCGATGACCTTGC 740  
QY 375 aSerArgGluTrpLeuProLeuAsnArgSerValAsnAsnValValValArgTyrGlyHi 395  
Db 741 TTTAGGAGGTGGCTTCCACTTAAACCGTTCTGTGAACAATGTGGTTGTAGATATGTTCA 800  
QY 395 sSerLeuAlaLeuTyrLysAspLysIleTyrMetTyrGlyLysIleAspPro-ThrG 415  
Db 801 TTTCTTGGCATATACAGGATAAAATTTACATGATGAGGAGAAATTTGATTC-ACTG 859  
QY 415 lyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuT 435  
Db 860 GGAATGTGACCAATGAGTGTGAGAGTTTTCACATTCATATGAGTCTCATGGTGTGTTGA 919  
QY 435 hrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuL 455  
Db 920 CCCCTAAGGCAAGGAGCAGTATGAGTGGTGGGCACTCTGCACACATTTGTACACTGA 979



QY	948	luCysMetTrrpCysSerAsnVetLysGlnCysValAspSerAsnAlaTyrValAlaSerP	968
Db	2912	AATGCATGTGGTGCAGCAACATGAAGCAGTGCCTAGACTCCAATGCCTAGCTAGCCTCCT	2971
QY	968	heProPheGlyGlnCysMetGluTrrpTyrThrMetSerThr-CysProProGluAsnCys	987
Db	2972	TTCTCTTCGGCCAGTGTATCGAGTGGTATACAAATGAGCAG-CTGCCACACCTGAAATATGC	3030
QY	988	SezGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrrpCysThrAsp	1007
Db	3031	TCTGGCTACTGTACTTCAGCCATTGCTTGGAGCAGCCAGGCTGTGGTGTGCTACTGAT	3090
QY	1008	ProSerAnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyProValLysMet	1027
Db	3091	CCATAGCAATACTGGGAAGCAATGATTTATGAGGAAGCTATTAAGACCCCGTAAAGATG	3150
QY	1028	ProSerGlnAlaPro--ThrGlyAsnPhe-Tyr-ProGlnProLeuLeuAsnSerSerMet	1046
Db	3151	CCTTCACA-TG-CCTCTACAGGAAATGT-GTACCACAGCCCTTCTGAACTCCAGCATG	3207
QY	1047	CysLeuGluAspSerArgTyrAsnTrrpSerPheIleHisCysProAlaCysGlnCysAsn	1066
Db	3208	TGTCATAGACACAGCATCAACTGTGCTTTTATTACATGTGCAGCTTCCCAAGTGCAT	3267
QY	1067	GlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn-LeuThrThrGln	1086
Db	3268	GGACACAGCAAGTGCATCAACAGAGCATCTGTGAGAAGTGTAGGA-CCTGACCACAGG	3326
QY	1086	LYseHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysCys	1106
Db	3327	CAAGCATCGAGACCTGCATATCTGGCTTCTATGTGTGACCGCACTAATGGAGGCAATG	3386
QY	1106	sGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPhe	1126
Db	3387	TCAGCCATGCAAGTGCAACGGCACGCATCACTGTGCAACACCAACACTGGCAAGTGCTT	3446
QY	1126	eCysThrThrLysGlyValLysGlyAsp-GluCysGlnLeuCysGluValGluAsnArgT	1146
Db	3447	CTGTACCACCAAGGGCGTCAAGGAGGA-GGAGTCCCAAGCTATGTGAGTGAAGAAATCGAT	3505
QY	1146	YrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPhe	1165
Db	3506	ACCAGGGAACCCCTCTCA-AGGAACATGTTACTATATACCTTCTCATGTACTATCAGTTT	3564
QY	1166	ThrPheSerLeuSerGlnLysAspAspArgTyrTyrThrAlaIleAsnPheValAlaThr	1185
Db	3565	ACTTTTAGCTGTCCCAAGGAGACGCGCTATTACACAGCCATCAATTTTGTGGCTACT	3624
QY	1186	ProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeu	1205
Db	3625	CCTGATGNACAAACACAGGATTTGGACATGTTTCATCAATGCCCTCCAAAACCTTCAACCTC	3684
QY	1206	AsnIleThrTrpAlaAla-SerPheSer-AlaGlyThrGlnAla-GlyGluGluMet-Pr	1224
Db	3685	AACATCACCTGGGCCAC-CAGTTTCCC-AGCTGGAACCCAGAC-TGAGAGAGAGGT-GCC	3740
QY	1224	oValValSerLysThrAsnIleLysGluTyrLYsAspSerPheSerAsnGluLysPheAs	1244
Db	3741	TGTTGTTTCAAAAACCAACATCAGGAATACAAAGATAGCTTCTCTAATGAGAAATTTGA	3800
QY	1244	pPheArgAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTrrpProil	1264
Db	3801	TTTTTCGAACCATCCAAACATTACTTTCTTTTGTATTATGTCAAGTAATTTCACTTGGCCCAT	3860
QY	1264	eLYsIleGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheVa	1284
Db	3861	CAAAATTCAAATTCGCTTCTCCAGCACAGCAATTTTATGACCTGGTGCACTTCTTCGT	3920
QY	1284	lThrPhePheSerCysPheLeuSerLeuLeuValAlaAlaValValTrrpLysIleLY	1304
Db	3921	GACTTTTCTTCAGCTGTTTCCCTCTCTCTGCTTCTGGTGGCGAGTGGTCTCGAAGATCAA	3980

Qy	1304	sGlnSerCysTrpAlaSerArgArgGluGlnLeuLeuArgGluMetGlnGlnMetAl	1322
Db	3981	GCAGAGCTGTGTGGCATCCAGGAGGAGAGCAACTTCTTCGGGAGATGCAACAGATGCC	4040
Qy	1324	aSerArgProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLe	1344
Db	4041	CAGCCGCCCTTTGCTTCTGTAAATGTTGCTTGGAAACAGATGAAGAACTCTCTGATCT	4100
Qy	1344	uileGlyGlySerIleLeysThrValProLysProIleAlaLeuGluProCysPheGlyAs	1364
Db	4101	CATTGGGGGAAGTATAAGACTGTCTCTAAGCCCATTCGCCCTCGAGGCCCTGCTTTGGTAA	4160
Qy	1364	nLysAlaAlaValLeuSerValPheValArgLeuProArgGlyLeuGlyGlyIleProPr	1384
Db	4161	CRAAGCTGCTGTCCTCTCTGTGTTTGAGAGCTCCCTCGAGAGCTGGAGAGATCCCTCC	4220
Qy	1384	oProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspIleSerGlnGlnMetPr	1404
Db	4221	TCTGGTCAGTCAGGTCCTGCTGGGCCAGTGCCCTGGTGGACATTTCTCAGCAGATGCC	4280
Qy	1404	oIleValTyrIlysGluIlysserGlyAlaValArgAsnArgIlysGlnGlnProProAlaGl	1424
Db	4281	AATAGTGTACAAGGAGAAGTCAGAGCTGTAAAGAACCGGAAGCAGAGCCCTTCGACA	4340
Qy	1424	nProGlyThrCysIle 1429	
Db	4341	GCCTGGAACTGCATT 4356	
RESULT 14			
AAZ91916			
ID	AAZ91916	standard; cDNA; 8827 BP.	
XX	AC	AAZ91916;	
XX	AC	AAZ91916;	
XX	AC	AAZ91916;	
DT	08-JUN-2000	(first entry)	
XX	DE	Wild type (C57BL/6J) murine mahogany protein coding sequence.	
XX	DE	Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;	
KW	KW	weight regulation; cell therapy; body weight disorder; cachexia;	
KW	KW	anorexia; hyperpigmentation; increased metabolic rate disorder;	
KW	KW	hyperphagia; Anticobesity; antianorexic; anticachexic; ss.	
XX	OS	Mus sp.	
XX	XX	WO2000005373-A2.	
XX	XX	03-FEB-2000.	
XX	XX	21-JUL-1999; 99WO-US016484.	
XX	XX	21-JUL-1998; 98US-0093630P.	
PR	PR	20-OCT-1998; 98US-0104978P.	
PR	PR	05-FEB-1999; 99US-00245041.	
XX	XX	(MILL-) MILLENIUM PHARM INC.	
PA	PI	Moore K, Nagle DL;	
XX	PI	WPI; 2000-195103/17.	
DR	DR	P-PSDB; AAY81803.	
XX	XX	New human and murine mahogany genes, useful, e.g. for diagnosis and	
PT	PT	treatment of body weight disorders.	
XX	XX	Claim 1; Fig 2a; 188pp; English.	
XX	XX	This sequence represents a murine mahogany gene of the invention. The	
CC	CC	mahogany genes are used: (i) to produce recombinant mahogany (mg)	
CC	CC	proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming	
CC	CC	therapeutics; (iii) as a source of diagnostic probes and primers for	
CC	CC	detecting expression of mg genes or mutations, regulatory defects, in	
CC	CC	this gene, or for isolation of related sequences; and (iv) in (cell-	

CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (II) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (II) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (II)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX  
 SQ Sequence 8827 BP; 2304 A; 2077 C; 2162 G; 2284 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 2,79e-149 Length: 8827  
 Score: 181.00 Matches: 575  
 Percent Similarity: 95.67% Conservative: 0  
 Best Local Similarity: 95.67% Mismatches: 13  
 Query Match: 12.67% Indels: 26  
 DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x AA291916 (1-8827)

CC	842	MetGlnSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys	861
CC	2790	ATGCAATCATCTCAAGATATGTCACAGCTCACCTGACTCCATGGTGTGCTTCGGAAG	2849
CC	862	IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln	881
CC	2850	ATCAATGTGCTTACTGGTCTCGGAGGATATCTCCATTCACAATAGTTTGTGTCAG	2909
CC	882	TrpMetProSerGluProSerAspAlaGlyPheCysGlyLeuSerGluProSerThr	901
CC	2910	TGGATGCCATCTGAGCCCACTGATGGCTCTGTGGGATCTTGTCAGAGCTAGTACT	2969
CC	902	ArgGlyLeuLysAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro	921
CC	2970	CGGGGATTAAGGCTGCAACCTGCATCAACCTCTCAATGGCAGCGCTGTGAAGGCT	3029
CC	922	AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp	941
CC	3030	GCAACCAACAGTCCCAAGCAGTCCCGGACACCAATGTCCTCGGACACGGTGTGGCA	3088
CC	942	CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspS	961
CC	3089	GTGCACCTAGCAG-CAGCTCGAGTGCATGTGTGTGCAGTAACATGAAGCAGTGTGTG	3147
CC	961	erAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerT	981
CC	3148	CCAATGCTTACGTGGGCTCTCCCTTTGGCCAGTGTATGGAATGGTATACATGAGCA	3207
CC	981	hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro	1000
CC	3208	G-CTGCCACCTCAAAATGCTCTGGCTACTGTACCTGACGAGCCATGCTTGAGCAGCA	3266
CC	1001	GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer	1020
CC	3267	GGCTGTGGTGGTGTACTATCTATCTAGCAATCTGGGAAAGGAAATGTATTGAGGCG	3326
CC	1021	TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln	1039
CC	3327	TATAAGGACCTCTGAAGATGCCCTCACAGGCTC-TG-CAGGAAATGT-GTATCCACAG	3383
CC	1040	ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis	1059
CC	3384	CCCTTTCTGAACCTCCAGCATGTCTCTAGAGGACAGCAGATACAACCTGGTCTTCTCA	3443
CC	1060	CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys	1079
CC	3444	TGTCCAGCTTGCCAGTGCACACGACACAGCAATGATCAACAGAGTATCTGTGAGAG	3503

QY	1080	CysGluAsn-LeuThrThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs	1099
DB	3504	TGTGAGGA-CTTGACACCGGSCAGACTGGAGACCTGTATATCTGGCTTCTATGTGA	3562
QY	1099	pProThrAnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerIleuCysAs	1119
DB	3563	CCCGACTAATGGAGCAAAATGTACAGCATGCAAGTGAATGGGCACGCATCATCTGTCAA	3622
QY	1119	nThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe	1139
DB	3623	CACCAACACCGGCAAGTCTTCTGTACCAACAAAGGTGTCAAGGGGACAGTCCAGCT	3682
QY	1139	uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL	1159
DB	3683	ATGTGAGGTAGAAAATCGATACCAAGGAACCTCTCAA-AGGAACATGCTACTATACCC	3741
QY	1159	euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA	1179
DB	3742	TTCTCATTTGACTATCAGTTACCTTTAGCTGTCTCCAGGAGACGACCGCTACTACAG	3801
QY	1179	IaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA	1199
DB	3802	CCATCACTTTGTGCTACTCTCTGATGAACAAACAGGGATTTGGACATGTTTCATCAATG	3861
QY	1199	IaSerLysAsnPheAsnLeuAsnIleThrTrpAlaIaIa-SerPheSer-AlaGlyThrGl	1218
DB	3862	CTCTCAAAATACTTCAACCTCAACATCACCTGGGCGCAC-CAGCTTCCC-AGCCGGAA	3919
QY	1218	nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer	1237
DB	3920	GAC-TGGAGAGAGAGT-GCCTGTGTGTTTCAAAACCAACATCAAGGAATCAAGATAGC	3977
QY	1238	PheSerAsnGluLysPheAspPheArgAsnHis-ProAsnIleThrPhePheValTyrVa	1257
DB	3978	TTCTCTAATAGAAATTTGATTTTTCGCAACGA-TCCAAACATCACTTCTTTTGTATGT	4036
QY	1257	IserAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMe	1277
DB	4037	CAGTAATTTCACTTGGGCCCATCAAAATTCAGATTGCTTCTCCAGCACAGCACTTCAT	4096
QY	1277	tAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAl	1297
DB	4097	GGACTGTGTACAGTCTTCTGTGACTTCTTTCAGTTGTTTCTCTCGCTGCTTCTGTGGC	4156
QY	1297	aAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluGlnLeuLe	1317
DB	4157	TGCAGTGTCTGGAAGATCAAGCAGAGCTGTGGGCATCCAGGCGGAGAGCAACTTCT	4216
QY	1317	uArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluTh	1337
DB	4217	TCGGAGATGCAACAGATGCGCCAGCCGCTTGTCTTCTGTAAACGTTGCCITGGAAAC	4276
QY	1337	rAspGluGluProProAspLeuIleGlySerIleLysThrValProLysProIleAl	1357
DB	4277	AGATGAGAGGCTCTCTGATCTTATTGGGGGAGTATTAAGACTGTTCCCAAAACCCATTGC	4336
QY	1357	aLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProAr	1377
DB	4337	ACTGAGCCGCTGTTTGGCAACAAGCCGCTGCTCTCTGTGTGTGTGAGGCTCCCTCG	4396
QY	1377	gGlyLeuGlyLysIleProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVa	1397
DB	4397	AGGCTGGGTGGCATCCCTCTCTGCGCAGTCAGGCTTGTCTGTGCTGTGAGGCTCCCTGT	4456
QY	1397	IaPheSerGlnGlnMetProIleValTyrLysGluLysSerGlyAlaValArgAsnAr	1417
DB	4457	GGACATTTCTCAGCAGATGCCGATAGTGTACAGGAGAGTTCAGGAGCCGTGAGAACCG	4516
QY	1417	gLyGlnGlnProProAlaGlnProGlyThrCysIle	1429
DB	4517	GAACGACGAGCCCTGCACAGCTGGGACCTGCATC	4553

AA578781  
ID AAS78781 standard; cDNA; 3165 BP.  
AC AAS78781;  
XX  
DT 13-FEB-2002 (first entry)  
DE DNA encoding novel human diagnostic protein #14585.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
FN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR P-PSDB; ABG14594.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 14585; 103pp; English.  
XX  
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activities. The  
CC polypeptide and polynucleotide sequences have application in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 3165 BP; 1042 A; 705 C; 587 G; 731 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 3165 Length: 3165  
Score: 142.00 Matches: 142  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 9.94% Indels: 0  
DB: 5 Gaps: 0

US-09-787-097-12 (1-1429) x AAS78781 (1-3165)

Qy 920 ArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939  
|||||

Db 1219 AGGCTGCAAAACACAGTGTAAAGCAGTGGCGACACCATGTGCTTGAGCAGCATGT 1278  
Qy 940 GlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959  
AC |||||  
Db 1279 GGAGATTGCACCGCGCAGCTCTGAGTGCATGTGTGTCAGCAACATGAAAGCAGTGTGTG 1338  
DT |||||  
Qy 960 AspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrThrMet 979  
DE |||||  
Db 1339 GACTCCAAATGCTATGTGGCTCTCTCCCTTTGGCCAGTGTATGGAAATGTTATACGATG 1398  
XX |||||  
Qy 980 SerThrCysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGln 999  
KW |||||  
Db 1399 AGCACCTGCCCTCCCTGAAATTTGTTCAAGCTACTGTACCTGTAGTCACTTCTTGAGCAA 1458  
XX |||||  
Qy 1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyLysCysLleGluGly 1019  
FN |||||  
Db 1459 CCAGGCTGTGGTGGTGTACTGATCCAGCAATACTGCGAAAGGAAATGCAATAGAGGGT 1518  
XX |||||  
Qy 1020 SerTyrLysGlyProValLysMetProSerGluAlaProThrGlyAsnPheTyrProGln 1039  
DE |||||  
Db 1519 TCCTATTAAGGACCGAGTGAAGATGCTTCGACAGCCCTACAGGAATTTCTATCCACAG 1578  
XX |||||  
Qy 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059  
KW |||||  
Db 1579 CCCCTGCTCAATCCAGCATGTGTCTAGAGGACAGCAGATACAACTGCTTTTCATTAC 1638  
XX |||||  
Qy 1060 CysPro 1061  
XX |||||  
Db 1639 TGTCCA 1644  
XX |||||  
RESULT 16  
AAS76294  
ID AAS76294 standard; cDNA; 3313 BP.  
XX  
AC AAS76294;  
XX  
DT 13-FEB-2002 (first entry)  
XX  
DE DNA encoding novel human diagnostic protein #12098.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US008631.  
XX  
PR 31-MAR-2000; 2000US-00540217.  
PR 23-AUG-2000; 2000US-00649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.  
DR P-PSDB; ABG12107.  
XX  
XX New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity.  
XX  
PS Claim 1; SEQ ID NO 12098; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)  
CC sequences. (I) is useful as hybridisation probes, polymerase chain  
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
CC and in recombinant production of (II). The polynucleotides are also used  
CC in diagnostics as expressed sequence tags for identifying expressed  
CC genes. (I) is useful in gene therapy techniques to restore normal  
CC activity of (II) or to treat disease states involving (II). (II) is  
CC useful for generating antibodies against it, detecting or quantitating a  
CC polypeptide in tissue, as molecular weight markers and as a food  
CC supplement. (II) and its binding partners are useful in medical imaging  
CC of sites expressing (II). (I) and (II) are useful for treating disorders  
CC involving aberrant protein expression or biological activities. The  
CC polypeptide and polynucleotide sequences have application in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
CC coding sequences of the invention. Note: The sequence data for this  
CC patent did not appear in the printed specification, but was obtained in  
CC electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 3165 BP; 1042 A; 705 C; 587 G; 731 T; 0 U; 0 Other;

CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC responsible for genetic disorders or other traits to assess biodiversity  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 3313 BP; 1090 A; 731 C; 724 G; 768 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 3,64e-115 Length: 3313  
 Score: 142.00 Matches: 142  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 9.94% Indels: 0  
 DB: 5 Gaps: 0

US-09-787-097-12 (1-1429) x AAS76294 (1-3313)

QY 920 ArgProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCys 939  
 DB 1219 AGCCCTCAACACACAGTGTAGCAGTCCCGACACCATGTCCTTGAGCAGCATGT 1278  
 QY 940 GlyAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysVal 959  
 DB 1279 GGAGATTGCAACACGCGCAGCTGTAGTGCATGTGTGCGAGCAACATGAAGCAGTGTGTG 1338  
 QY 960 AspSerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTrpTrpMet 979  
 DB 1339 GATCCAAATGCCATGTGGCTCTCTCCCTTTGGCCAGTGATGGAATGGTATACGATG 1398  
 QY 980 SerThrCysProGluAsnCysSerGlyTrpCysThrCysSerHisCysLeuGluGln 999  
 DB 1399 AGCACCTGCCCTCGAAATTTCTCAGCTACTGTACCTGTAGTCATTGCTTGGAGCAA 1458  
 QY 1000 ProGlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyCysLysGluGly 1019  
 DB 1459 CCAGGCTGTGGCTGTACTGATCCAGCAATPACTGGCAAGGGAATGCCATAGAGGT 1518  
 QY 1020 SerTrpLysGlyProValLysMetProSerGlnAlaProThrGlyAsnPheTrpProGln 1039  
 DB 1519 TCCTATAGACACAGTGAATGCTTCCAGGCCCTACAGGAATTTCTATCCACAG 1578  
 QY 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTrpAsnTrpSerPheLeHis 1059  
 DB 1579 CCCTGTCTCAATTCAGCATGTCTAGAGGACAGCAGATACAACTGGTCTTTTCATTCAC 1638  
 QY 1060 Cyspro 1061  
 DB 1639 TGCCA 1644

RESULT 17

AAS72658  
 ID AAS72658 standard; cDNA; 671 BP.

XX AC AAS72658;  
 XX AC AAS72658;  
 DT 13-FEB-2002 (first entry)

XX DNA encoding novel human diagnostic protein #8462.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;  
 KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.  
 OS WO200175067-A2.  
 FN 11-OCT-2001.  
 PD 30-MAR-2001; 2001WO-US008631.  
 XX 31-MAR-2000; 2000US-00540217.  
 PR 23-AUG-2000; 2000US-00649167.  
 XX (HYSB-) HYSSEQ INC.  
 PI Drmanac RT, Liu C, Tang YT;  
 XX WPI; 2001-639362/73.  
 DR P-FSDB; A8G08471.  
 XX New isolated polynucleotide and encoded polypeptides, useful in  
 PT diagnostics, forensics, gene mapping, identification of mutations  
 PT responsible for genetic disorders or other traits and to assess  
 PT biodiversity.  
 XX Claim 1; SEQ ID NO 8462; 103pp; English.  
 PS The invention relates to isolated polynucleotide (I) and polypeptide (II)  
 CC sequences. (I) is useful as hybridisation probes, polymerase chain  
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,  
 CC and in recombinant production of (II). The polynucleotides are also used  
 CC in diagnostics as expressed sequence tags for identifying expressed  
 CC genes. (I) is useful in gene therapy techniques to restore normal  
 CC activity of (II) or to treat disease states involving (II). (II) is  
 CC useful for generating antibodies against it, detecting or quantitating a  
 CC polypeptide in tissue, as molecular weight markers and as a food  
 CC supplement. (II) and its binding partners are useful in medical imaging  
 CC of sites expressing (II). (I) and (II) are useful for treating disorders  
 CC involving aberrant protein expression or biological activity. The  
 CC polypeptide and polynucleotide sequences have applications in  
 CC diagnostics, forensics, gene mapping, identification of mutations  
 CC and to produce other types of data and products dependent on DNA and  
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic  
 CC coding sequences of the invention. Note: The sequence data for this  
 CC patent did not appear in the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 671 BP; 207 A; 147 C; 148 G; 169 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1,72e-87 Length: 671  
 Score: 110.00 Matches: 110  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 7.70% Indels: 0  
 DB: 5 Gaps: 0

US-09-787-097-12 (1-1429) x AAS72658 (1-671)

QY 1158 ThrLeuLeuLeuAspTrpGlnPheThrPheSerLeuSerGlnGluAspAspArgTrpTyr 1177  
 DB 125 ACTCTTCTATTGACTATCAGTTTCACCTTCAGTCTATCCCGAGGAGATGATCCTATTAC 184  
 QY 1178 ThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIle 1197  
 DB 185 ACAGCTATCATTTTGGTCTACTCTCTGACGACAAACAGGATTTGGACATGTTTCA 244  
 QY 1198 AsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaIleAsnSerAlaGlyThr 1217  
 DB 245 AATGCCTCCAGAAATTTCAACCTCAACATCAGCTGGGCTGCCAGTTTCTCAGCTGAACC 304  
 QY 1218 GlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysGluTrpLysAspSer 1237



Db 305 CAGGCTGGAGAGAGATGCTGTTCTTCAAAACCAACATTAAAGGAGTACAAAGATAGT 364  
Qy 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTrVal 1257  
Db 365 TTCTCTTAAGAGAAGTTGATTTTCGCAACCAACCAATATACATTTCTTTGTTTAAGTC 424  
Qy 1258 SerAsnPheThrTrpProIleLysIleGln 1267  
Db 425 AGTAATTTCACTGGGCCCATCAAAATTCAG 454

## RESULT 18

AAS02409  
ID AAS02409 standard; cDNA; 668 BP.

AC AAS02409;

DT 18-JUL-2001 (first entry)

XX Human secreted protein, cDNA #15.

XX Human; secreted protein; immunogen; antibody; diagnosis;  
KW rheumatoid arthritis; hyperproliferative disorder; neoplasm;  
KW cardiovascular disorder; cerebrovascular disorder; cerebral ischaemia;  
KW angiogenesis; Alzheimer's disease; bacterial infection; viral infection;  
KW fungal infection; corneal infection; wound healing; cell culture;  
KW epithelial cell proliferation; skin ageing; transplantation;  
KW tissue regeneration; chemotaxis; food additive; ss.

XX Homo sapiens.

XX WO200123546-A1.

XX 05-APR-2001.

XX 26-SEP-2000; 2000WO-US026323.

XX 27-SEP-1999; 99US-0155805P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben S, Komatsoulis GA;

XX WPI: 2001-266150/27.

XX P-PSDB; AAU01736, AAU01774.

XX Nucleic acids encoding 37 human secreted polypeptides, useful for  
PT preventing, diagnosing and/or treating e.g. cancers, Parkinson's disease  
PT and diabetic retinopathy.

XX Disclosure; Page 420; 494pp; English.

XX The sequence encodes a human secreted protein of the invention. The  
CC polynucleotides, polypeptides and antibodies raised against them are used  
CC to prevent, treat or ameliorate a medical condition in e.g. humans, mice,  
CC rabbits, goats, horses, cats, dogs, chickens or sheep. The  
CC polynucleotides and antibodies are also used in diagnosing a pathological  
CC condition or susceptibility to a pathological condition. The antibodies  
CC can also be used in alleviating symptoms associated with the disorders  
CC and in diagnostic immunoassays e.g. radioimmunoassays or enzyme linked  
CC immunoassay assays (ELISA). Disorders which are diagnosed or treated  
CC include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative  
CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders  
CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,  
CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,  
CC infections caused by bacteria, viruses and fungi and ocular disorders  
CC e.g. corneal infection. The polypeptides can also be used to aid wound  
CC healing and epithelial cell proliferation, to prevent skin aging due to  
CC sunburn, to maintain organs before transplantation, for supporting cell  
CC culture of primary tissues, to regenerate tissues and in chemotaxis. The  
CC polypeptides can also be used as a food additive or preservative to  
CC increase or decrease storage capabilities

SQ Sequence 668 BP; 298 A; 113 C; 120 G; 137 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 3,48e-82 Length: 668  
Score: 104.00 Matches: 104  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 7.28% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAS02409 (1-668)

Qy 1164 GlnPheThrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheVal 1183  
Db 11 CAGTTTCACTTTAGTCTATCCAGGAAGATGCGCTATTACACGCTATCAATTTGTG 70

Qy 1184 AlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPhe 1203  
Db 71 GCTACTCTCGACGAACAAACAGGATTTGACATGTTTCATCAATGCTCCAAAGATTTTC 130

Qy 1204 AsnLeuAsnIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMet 1223  
Db 131 AACCTCAACATCACTGGGTGCCAGTTCTCTAGCTGGAACCCAGGCTGGAGAGAGATG 190

Qy 1224 ProValValSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPhe 1243  
Db 191 CCGTCTGTTTCAAAACCAACATTAAGAGTACAAAGATAGTTCTCTAATGAGAAGTTT 250

Qy 1244 AspPheArgAsnHisProAsnIleThrPhePheValTrValSerAsnPheThrTrpPro 1263  
Db 251 GATTTTCGCAACCAACCAATATCACTTTCTTTGTTTATGTCAGTAATTTCCACTGGCCCC 310

Qy 1264 IleLysIleGln 1267

Db 311 ATCAAAATTCAG 322

## RESULT 19

AAZ91917

ID AAZ91917 standard; cDNA; 2419 BP.

XX AC AAZ91917;

XX 08-JUN-2000 (first entry)

XX Murine mahogany protein coding sequence akml003.

XX Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
KW weight regulation; cell therapy; body weight disorder; cachexia;  
KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.

XX Mus sp.

XX WO200005373-A2.

XX 03-FEB-2000.

XX 21-JUL-1999; 99WO-US016484.

XX 21-JUL-1998; 98US-0093630P.

XX 20-OCT-1998; 98US-0104978P.

XX 05-FEB-1999; 99US-00245041.

XX (MILL-) MILLENIUM PHARM INC.

XX Moore K, Nagle DL;

XX WPI: 2000-195103/17.

XX P-PSDB; AAY81804.

XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
PT treatment of body weight disorders.

XX

PS Claim 1; Fig 8a; 188pp; English.

XX This sequence represents a murine mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (ii); (iii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
 CC identify other (extracellular) products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (ii) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (ii)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX  
 SQ Sequence 2419 BP; 578 A; 567 C; 680 G; 594 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 8,11e-73 Length: 2419  
 Score: 94.00 Matches: 94  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.58% Indels: 0  
 DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x AAZ91917 (1-2419)

Qy 513 AlaLeuTyrValHisGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAsp 532  
 Db 1803 GCTCTAGCTTCATGGTGGTACAGGCTTTTCAGGCCACAAATACCGGCTTCAGAT 1862  
 Qy 533 AspLeuTyrArgTyrAspValAlaThrGlnMetTyrThrIleLeuLysAspSerArgPhe 552  
 Db 1863 GACCTCTACAGATACCATGTGGATCTCAGATGTGGACCAATCTTAAAGACAGCGGATTT 1922  
 Qy 553 PheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572  
 Db 1923 TTCGGTACTTGCATACAGCTGTGATGAGTGGACCATCTGCTGTTTGGAGGGAAC 1982  
 Qy 573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheValAla 592  
 Db 1983 ACACAAATGACACTTCCATGAGCCACGGTGCCAAATGCTTCTCTCGGACTTCATGGCT 2042  
 Qy 593 TyrAspIleAlaCysAspArgTyrPheSerValLeuProArgPro 606  
 Db 2043 TATGACATTTGTTGACCGATGGTCAGTGTCTTCCAGACCT 2084

RESULT 20  
 AAL14686/c  
 ID AAL14686 standard; cdna; 481 BP.

XX AAL14686;  
 XX  
 XX  
 DT 07-DEC-2001 (first entry)

XX Human breast cancer expressed polynucleotide 7143.  
 DE  
 XX Human; breast cancer; cell marker; cytostatic; ss.

XX Homo sapiens.

XX WO200151628-A2.

XX 19-JUL-2001.

XX 10-JAN-2001; 2001WO-US000799.

PR 14-JAN-2000; 2000US-0176077P.  
 PR 14-MAR-2000; 2000US-0189167P.  
 PR 24-MAR-2000; 2000US-0132099P.  
 PR 29-MAR-2000; 2000US-0193480P.  
 PR 15-MAY-2000; 2000US-0205230P.  
 PR 09-JUN-2000; 2000US-0211315P.  
 PR 25-JUL-2000; 2000US-0220534P.  
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.  
 XX  
 PI Lillie J, Xu Y, Wang Y, Steinmann K;  
 XX  
 DR WPI; 2001-451856/48.  
 XX  
 PT New peptide useful as a marker for the diagnosis of breast cancer.  
 XX  
 PS Claim 1; Page 1292; 3695pp; English.  
 XX  
 CC The invention relates to human breast cancer expressed polynucleotides  
 CC (AAL07544-AAL26789) and methods of assessing whether a patient is  
 CC afflicted with breast cancer by examining the correlation between the  
 CC expression of certain markers and the cancerous state of breast cells.  
 CC The polynucleotides and encoded polypeptides are potential markers for  
 CC detecting, diagnosing, monitoring, characterising treating and  
 CC potentially preventing breast cancer. The polynucleotides and encoded  
 CC polypeptides are also useful for isolating compounds with cytostatic  
 CC activity  
 XX  
 SQ Sequence 481 BP; 106 A; 133 C; 126 G; 116 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1.07e-71 Length: 481  
 Score: 92.00 Matches: 92  
 Percent Similarity: 100.00% Conservativity: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 6.44% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAL14686 (1-481)

Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
 Db 277 ACTCGGGAGCTGAAGCTGCAACCTCATCAACCCACTCAATGTTAGTGTCTGTGAAGG 218  
 Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
 Db 217 CCTGCAAAACACACAGTGTAGCAGTCCGGACACCATGTGCTTGAGGACAGCATGTGA 158  
 Qy 941 AspCysThrSerGlySerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAsp 960  
 Db 157 GATTGCACACAGCGGACCTGTGAGTGTATGTGTGGCAGCAACATGAAGCAGTGTGGAC 98  
 Qy 961 SerAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrThrMetSer 980  
 Db 97 TCCAATGCCCTATGTGGCTCTCTCCCTTTTGGCCAGTGTATGAATGGTATAGATGAGC 38  
 Qy 981 ThrCysProProGluAsnCysSerGlyTyrCysThr 992  
 Db 37 ACCTGCCCCCTGAAAATTGTTTCAGGCTACTGTACC 2

RESULT 21  
 AAX84568  
 ID AAX84568 standard; cdna to mRNA; 1012 BP.

XX AAX84568;

XX  
 DT 13-SEP-1999 (first entry)

XX Biorhythm marker gene.

XX Biorhythm marker gene; circadian rhythm; ss.

XX Rattus sp.



XX ACAL11257;  
 XX AC (FANG//) BANGUR C S.  
 XX DT 05-JUN-2003 (first entry)  
 XX DE Human lung adenocarcinoma library cDNA SEQ ID 966.  
 XX DE Human; lung cancer; ss; lung tumour; cytostatic; vaccine;  
 XX DE T cell expansion; CD4; CD8.  
 XX KW Homo sapiens.  
 XX OS US2002197669-A1.  
 XX PN 26-DEC-2002.  
 XX PD 03-MAY-2001; 2001US-00849626.  
 XX PF 13-DEC-2000; 2000US-00736457.  
 XX PR (BANG//) BANGUR C S.  
 XX PA (FANG//) FANGER G R.  
 XX PA (WANG//) WANG A.  
 XX PA (WANG//) WANG T.  
 XX PA (SWIT//) SWITZER A P.  
 XX PA (MCNE//) MCNEILL P D.  
 XX PA (CLAP//) CLAPPER J D.  
 XX PI Bangur CS, Fanger GR, Wang A, Wang T, Switzer AP, Mcneill PD;  
 XX PI Clapper JD;  
 XX PI WPI; 2003-352750/33.  
 XX DR Novel lung cancer polynucleotide encoding lung cancer protein, useful for  
 XX PT detecting the presence of lung cancer in a patient, and in pharmaceutical  
 XX PT compositions, e.g. vaccines, for treating lung cancer.  
 XX PS Example 1; Page; 72pp; English.

XX The invention relates to a polynucleotide encoding a lung tumour protein,  
 CC comprising a sequence selected from any of the 14 sequences mentioned in  
 CC the specification, or a sequence (S2) mentioned in specification,  
 CC complement of S1, sequences consisting of at least 20 contiguous residues  
 CC of S1, sequences that hybridise to S1, sequences having 75%, preferably  
 CC 90%, identity to S1, or degenerate variants of S1. Also included are an  
 CC isolated polypeptide (comprising a sequence (S3) selected from any one of  
 CC the 4 amino acid sequences mentioned in the specification, a sequence  
 CC encoded by the polynucleotide, or sequences having at least 70%,  
 CC preferably 90%, identity to a sequence encoded by the polynucleotide), an  
 CC expression vector comprising the polynucleotide operably linked to an  
 CC expression control sequence, a host cell transformed or transfected with  
 CC the vector, an isolated antibody (or its antigen-binding fragment) that  
 CC specifically binds to the polypeptide, detecting the presence of a cancer  
 CC in a patient, a fusion protein comprising the polypeptide, an  
 CC oligonucleotide that hybridises to S1 under moderately stringent  
 CC conditions, stimulating and/or expanding T cells specific for a tumour  
 CC protein (comprising contacting T cells with the polynucleotide, protein  
 CC or antigen-presenting cells, under conditions and for a time sufficient  
 CC to permit the stimulation and/or expansion of T cells) and inhibiting the  
 CC development of a cancer in a patient (by incubating CD4<sup>+</sup> and/or CD8<sup>+</sup> T  
 CC cells isolated from a patient with the polynucleotide, protein or antigen  
 CC presenting cells that express the polynucleotide, such that T cells  
 CC proliferate, administering to the patient an effective amount of the  
 CC proliferated T cells, and thus inhibiting the development of a cancer in  
 CC the patient. The polynucleotide, protein and cells are useful in a  
 CC composition for stimulating an immune response in a patient, and for  
 CC treating a cancer in a patient (particularly lung cancer). The  
 CC oligonucleotide is useful for determining the presence of a cancer in a  
 CC patient. The protein and oligonucleotides are useful in pharmaceutical  
 CC compositions, e.g. vaccines. The polynucleotide is also useful as a probe  
 CC or primer for nucleic acid hybridisation, and in the design and  
 CC preparation of ribozyme molecules for inhibiting expression of tumour  
 CC polypeptides and proteins in tumour cells. An amplified portion of the

CC polynucleotide is useful for isolating a full-length gene from a suitable  
 CC library. The present sequence is a cDNA (full length, extended or  
 CC partial) isolated from a library derived from lung tumour/cancer cells.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from the  
 CC USPTO at [seqdata.uspto.gov/sequence.html?DocId=20020197669](http://seqdata.uspto.gov/sequence.html?DocId=20020197669)

XX SQ Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.:	4.04e-63	Length:	246
Score:	82.00	Matches:	82
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	5.74%	Indels:	0
DB:	7	Gaps:	0

US-09-787-097-12 (1-1429) x ACAL11257 (1-246)

QY	839	LeuArgileMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly	858
DB	246	CTGCGAATAATGCAATCATCTCAGAGCATGTCACCAAGCTCACCTTAACCCCATGGTGGC	187
QY	859	LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer	878
DB	186	CTTCGGAAGATCAATGTGTCTCTACTGTGTCTGGAGAGATATGCCCCAATTTACAAATAGT	127
QY	879	LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu	898
DB	126	TTACTACAGTGGATCGCGTCTGAGCCAGATGCTGGATTCTGTGGAATTTATCAGAA	67
QY	899	ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys	918
DB	66	CCCACTACTCGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTCTGT	7
QY	919	GluArg 920	
DB	6	GAAAGG 1.	

RESULT 24  
 ACA02443/c  
 ID ACA02443 standard; cDNA; 246 BP.  
 XX AC ACA02443;  
 XX DT 22-MAY-2003 (first entry)  
 XX DE Lung cancer therapy and diagnosis associated cDNA #932.  
 XX KW Lung cancer; cytostatic; vaccine; gene therapy; cancer; gene; ss.  
 XX OS Homo sapiens.  
 XX PN US2002172952-A1.  
 XX PD 21-NOV-2002.  
 XX PF 10-JUL-2001; 2001US-00902941.  
 XX PR 30-JUN-1999; 99US-00346492.  
 XX PR 15-OCT-1999; 99US-00419356.  
 XX PR 17-DEC-1999; 99US-00468667.  
 XX PR 30-DEC-1999; 99US-00476300.  
 XX PR 06-MAR-2000; 2000US-00519642.  
 XX PR 22-MAR-2000; 2000US-00533077.  
 XX PR 10-APR-2000; 2000US-00548259.  
 XX PR 27-APR-2000; 2000US-00560406.  
 XX PR 05-JUN-2000; 2000US-00589184.  
 XX PR 11-JUL-2000; 2000US-00614124.  
 XX PR 29-AUG-2000; 2000US-00651563.  
 XX PR 08-SEP-2000; 2000US-00658824.  
 XX PR 26-SEP-2000; 2000US-00671325.  
 XX PR 06-OCT-2000; 2000US-00677419.

```
PR 30-OCT-2000; 2000US-00702705.
PR 13-DEC-2000; 2000US-00736457.
PR 03-MAY-2001; 2001US-00849626.
XX (CORI-) CORIXA CORP.
XX Henderson RA, Wang T, Watanabe Y, Johnson JC, Retter MW;
PI Durham M, Carter D, Fanger GK, Vedvick TS, Bangur CS, McNabb A;
XX WPI; 2003-328427/31.
XX
XX New polynucleotide, useful for preparing a composition for treating or
PT inhibiting development of cancer, e.g. lung cancer.
XX
XX Example 1; SEQ ID NO 966; 82pp; English.
XX
XX The invention describes an isolated polynucleotide comprising one of 32
CC sequences, complement or degenerate variants of them. The polynucleotide
CC is useful for preparing a composition e.g. a vaccine or for gene therapy,
CC for treating or inhibiting development of cancer, e.g. lung cancer. This
CC sequence represents a polynucleotide associated with the compositions and
CC methods for the therapy and diagnosis of lung cancer
XX
XX Sequence 246 BP; 63 A; 61 C; 60 G; 62 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 4,04e-63 Length: 246
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.74% Indels: 0
DB: 7 Gaps: 0
US-09-787-097-12 (1-1429) x ACA02443 (1-246)
QY 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValgly 858
Db 246 CTCGGAATAATCAGTCATCTCAGAGCATGTCACAGTCACCTTAACCCCATGGTGGC 187
QY 859 LeuArgLysIleAsnValSerTrpCysTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTCGGAAGATCAATGTGTCCTACTGCTGGAGAGATATGTCCTCCCATTTACAATAGT 127
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACACTGGATGCGCTCGAGCCAGTCATGCTGGATTCTGTGGAATTTATCAGAA 67
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTCTGT 7
QY 919 GluArg 920
Db 6 GAAGG 1
RESULT 25
AAL23548/c
ID AAL23548 standard; cDNA; 482 BP.
XX
XX AAL23548;
XX
XX 07-DEC-2001 (first entry)
XX
XX Human breast cancer expressed polynucleotide 16005.
XX
XX Human; breast cancer; cell marker; cytostatic; ss.
XX
XX Homo sapiens.
XX
XX WO200151628-A2.
XX
XX 19-JUL-2001.
XX
```

```
PF 10-JAN-2001; 2001WO-US000798.
XX
XX 14-JAN-2000; 2000US-0176077P.
PR 14-MAR-2000; 2000US-0189167P.
PR 24-MAR-2000; 2000US-0192099P.
PR 29-MAR-2000; 2000US-0193480P.
PR 15-MAY-2000; 2000US-0205230P.
PR 09-JUN-2000; 2000US-0211315P.
PR 25-JUL-2000; 2000US-0220534P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Lillie J, Xu Y, Wang Y, Steinmann K;
PI
XX WPI; 2001-451856/48.
XX
XX New peptide useful as a marker for the diagnosis of breast cancer.
PT
XX Claim 1; Page 2922; 3695pp; English.
PS
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity
XX
XX Sequence 482 BP; 106 A; 133 C; 126 G; 117 T; 0 U; 0 Other;
SQ
Alignment Scores:
Pred. No.: 3,41e-60 Length: 482
Score: 79.00 Matches: 92
Percent Similarity: 98.92% Conservative: 0
Best Local Similarity: 98.92% Mismatches: 0
Query Match: 5.53% Indels: 1
DB: 4 Gaps: 0
US-09-787-097-12 (1-1429) x AAL23548 (1-482)
QY 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920
Db 278 ACTCGGGGACTGAAGGCTGCACCTGCATCAACCCACTCAATGGTAGTGTGTGAAGG 219
QY 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAla-CysG1 940
Db 218 CTTGCCAAACACACAGTGTAGCAGTGCCTGACACCATGTGCTTGAGGACAGCAATGG 159
QY 940 YAspCysThrSerGlySerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAs 960
Db 158 AGATTGCACACAGCGCAGCTCTGAGTGTATGTGTGTCAGCAACATGAAGCAGTGTGTGGA 99
QY 960 pSerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTrpTrpThrMetSe 980
Db 98 CTCCAAATGCCTATGTGGCCTCCCTTCCTTTGGCCAGTGTATGGAATGATACATGATG 39
QY 980 rThrCysProGluAsnCysSerGlyTrpCysThr 992
Db 38 CACCTGCCCTCTGAAAATTTGTTTCAGGCTACTGTACC 2
RESULT 26
ABZ72040
ID ABZ72040 standard; DNA; 207433 BP.
XX
XX AC ABZ72040;
XX
XX 03-APR-2003 (first entry)
XX
XX Gene 216 H194BAC1098L22 nucleotide sequence SEQ ID NO 5.
XX
XX Human; Gene 216; chromosome 20p13-p12; antiasthmatic; anorectic;
XX
```



Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 5.39% Indels: 0  
 DB: 7 Gaps: 0

US-09-787-097-12 (1-1429) x ABX74891 (1-207433)

QY 985 GluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlulnProGlyCysGlyTrp 1004  
 Db 9815 GAAATTTGTTACGGCTACTGTACCTAGTCTTGTGAGCAACCGAGCTGTGGCTGG 9874

QY 1005 CysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTyrLysGlyPro 1024  
 Db 9875 TGACTGATCCCGCAATACTGGCAAAGGGAATGCATAGAGGGTTCCTATTAAGACCA 9934

QY 1025 ValIysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSer 1044  
 Db 9935 GTGAAGATGCTTCGCAAGCCCTACAGGAATTTCTATCCACAGCCCTGCTCAATTCC 9994

QY 1045 SerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysPro 1061  
 Db 9995 AGCATGTGTCTAGAGACACGACATACAACTGCTCTTTTCATTCACTGTCCA 10045

RESULT 28  
 AAZ91918  
 ID AAZ91918 standard; cDNA; 1051 BP.  
 XX  
 AC AAZ91918;  
 XX  
 DT 08-JUN-2000 (first entry)  
 XX  
 DE Murine mahogany protein coding sequence akml1004.  
 XX  
 KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antiobesity; antianorexic; anticachexic; ss.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200005373-A2.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 21-JUL-1999; 99WO-US016484.  
 XX  
 PR 21-JUL-1998; 98US-0093630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 99US-00245041.  
 XX  
 PA (MILL-) MILLENIUM PHARM INC.  
 XX  
 PI Moore K, Nagle DL;  
 XX  
 DR WPI; 2000-195103/17.  
 DR P-PSDB; AAY81805.  
 XX  
 PT New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.  
 XX  
 PS Claim 1; Fig 9a; 188pp; English.  
 XX  
 CC This sequence represents a murine mahogany gene of the invention. The  
 CC mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (ii); (ii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (ii) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (ii) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (ii) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (ii)-expressing cells intended for cell therapy, and as

CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity. (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX  
 SQ Sequence 1051 BP; 306 A; 243 C; 313 G; 189 T; 0 U; 0 Other;

Alignment Scores: 3.17e-57 Length: 1051  
 Pred. No.: 76.00 Matches: 111  
 Score: 98.23% Conservative: 0  
 Percent Similarity: 98.23% Mismatches: 1  
 Best Local Similarity: 5.32% Indels: 2  
 Query Match: 3 Gaps: 0  
 DB: 0

US-09-787-097-12 (1-1429) x AAZ91918 (1-1051)

QY 87 AlalaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
 Db 463 CGCGCGCGGTGTCGGCTCGGCGCGAGCCAGCAAGGAATGTACCGCGCGGTGTC 522

QY 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys-ProAlaGlyTrpValG1 126  
 Db 523 AACGGCGCGCGCTGCAACCCCTGGCCAGCGCGGTGCTGCCCCACG-GGCTGGGTGGG 581

QY 126 YGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValTh 146  
 Db 582 CGAGCAATGCCAGCACTGGCGGGCGCGCTTCAGACTAACTGGCTCTTCTGGATTGTAAC 641

QY 146 rAspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeulleGluGlyGlnPr 166  
 Db 642 AGATGGACCTGGGAATTATATAATAGACCAAGTGCACATGGCTCATTTGAAGACAGCC 701

QY 166 oAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLe 186  
 Db 702 AAATAGATATATGAGACTTCGCTTCACCACTTTTGTACAGATGTAGTGGGACCAATT 761

QY 186 uTyrValTyrAspGlyAspSerIleTyrAlaProLeu 198  
 Db 762 ATATGTTTATGATGGGACTCAATCTACGCACTCTG 798

RESULT 29  
 AAZ91925  
 ID AAZ91925 standard; DNA; 90050 BP.  
 XX  
 AC AAZ91925;  
 XX  
 DT 08-JUN-2000 (first entry)  
 XX  
 DE Wild type (C57BL/6J) murine mahogany protein genomic sequence c110/111.  
 XX  
 KW Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antiobesity; antianorexic; anticachexic; ds.  
 XX  
 OS Mus sp.  
 XX  
 PN WO200005373-A2.  
 XX  
 PD 03-FEB-2000.  
 XX  
 PF 21-JUL-1999; 99WO-US016484.  
 XX  
 PR 21-JUL-1998; 98US-0093630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 99US-00245041.  
 XX  
 PA (MILL-) MILLENIUM PHARM INC.  
 XX  
 PI Moore K, Nagle DL;





KW primary ciliary dyskinesia; pulmonary hypertension;  
KW hyaline membrane disease.

OS Homo sapiens.

XX WO200186003-A2.

PN 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

DR Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples.

XX Claim 1; SEQ ID NO 11514; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon  
CC nucleic acid probes for measuring gene expression in a sample derived  
CC from human lung comprising single exon nucleic acid probes having one of  
CC 12614 nucleic acid sequences mentioned in the specification, or their  
CC complements or the 12387 open reading frames derived from the 12614  
CC probes. Also included are a microarray comprising the novel set of probes  
CC : the novel set of probes which hybridize at high stringency to a nucleic  
CC acid expressed in the human lung; measuring gene expression in a sample  
CC derived from human lung, comprising (a) contacting the array with a  
CC collection of detectably labeled nucleic acids derived from human lung  
CC mRNA, and (b) measuring the label detectably bound to each probe of the  
CC array; identifying exons in a eukaryotic genome, comprising (a)  
CC algorithmically predicting at least one exon from genomic sequences of  
CC the eukaryote; and (b) detecting specific hybridisation of detectably  
CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
CC having a fragment identical to the predicted exon, the probe is included  
CC in the above mentioned microarray; assigning exons to a single gene,  
CC comprising (a) identifying exons from genomic sequence by the method  
CC above and (b) measuring the expression of each of the exons in several  
CC tissues and/or cell types using hybridisation to a single exon  
CC microarrays having a probe with the exon, where a common pattern of  
CC expression of the exons in the tissues and/or cell types indicates that  
CC the exons should be assigned to a single gene; a peptide comprising one  
CC of 12011 sequences, mentioned in the specification, or encoded by the  
CC probes/open reading frames (ORF). The probes are used for gene expression  
CC analysis, and for identifying exons in a gene, particularly using human  
CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
CC Pudlak syndrome, sarcoidosis, pulmonary haemorrhoidosis, pulmonary  
CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
CC present sequence is a single exon probe of the invention. Note: The  
CC sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 531 BP; 145 A; 82 C; 107 G; 197 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2.63e-51 Length: 531

Score: 69.00 Matches: 69  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.83% Indels: 0  
DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x ABS11523 (1-531)

QY 414 ThrGlyAsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTTPValLeu 433

Db 186 ACTGGGAATGTGACCAATGAGTTGAGAGTTTTCACATTCATTAATGAGTCATGGGTGTTG 245

QY 434 LeuThrProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThr 453

Db 246 TTGACCCCTTAAGCAAGAGGAGCAGTATGAGTGGTGGGCACTCTGCACACATTTGTACA 305

QY 454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473

Db 306 CTGAAGAATGGCCAGTGGTCACTGGTTCATCTTTGGTCACTGCCCTCTCTATGATAT 365

QY 474 IleSerAsnValGlnGlnTyrAspLeu 482

Db 366 ATAAGCAATGTGCAGGAATATGATTG 392

RESULT 32

AAI56405

ID AAI56405 standard; DNA; 193 BP.

XX AAI56405;

AC AAI56405;

DT 17-OCT-2001 (first entry)

DE Probe #25091 used to measure gene expression in human placenta sample.

XX Probe; microarray; human; placenta; antenatal diagnosis;

KW genetic disorder; ss.

XX Homo sapiens.

OS WO200157272-A2.

PN 09-AUG-2001.

PD 30-JAN-2001; 2001WO-US000663.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

DR Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human placenta.

PT Claim 25; SEQ ID NO 25091; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP).

CC The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders

XX SQ Sequence 193 BP; 50 A; 36 C; 52 G; 55 T; 0 U; 0 Other;

Alignment Scores:

```

Pred. No.: 2,74e-47 Length: 193
Score: 64.00 Matches: 64
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.48% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AA156405 (1-193)
QY 416 AsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThr 435
DB 2 AATGTGACCAATGAGTGTGAGATTTTCATTCATTAATGAGTCAGGGTGTTGTGACC 61
QY 436 ProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLys 455
DB 62 CCTAAGGCAAGGAGCAGATGATGAGTGTGGGCACTCTGCACACATTGTACACTGAAG 121
QY 456 AsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSer 475
DB 122 AATGGCCGAGTGTGATGCTGTGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAGC 181
QY 476 AsnValGlnGlu 479
DB 182 AATGTGCAGGAA 193

RESULT 33
ID ABS23881 standard; DNA; 193 BP.
XX AC ABS23881;
XX DT 19-AUG-2002 (first entry)
XX DE Human genome-derived single exon probe ORF from lung SEQ ID No 23872.
XX KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;
XX KW chronic obstructive pulmonary disease; interstitial lung disease;
XX KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
XX KW tuberculous scleroidosis; Gaucher's disease; Niemann-Pick disease;
XX KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
XX KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
XX KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
XX KW primary ciliary dyskinesia; pulmonary hypertension;
XX KW hyaline membrane disease; open reading frame; ORF.
XX OS Homo sapiens.
XX PN WO200186003-A2.
XX PD 15-NOV-2001.
XX PF 30-JAN-2001; 2001WO-US000665.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2002-114183/15.
XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.
XX PS Claim 4; SEQ ID NO 23872; 634pp; English.
XX DT 21-AUG-2000 (first entry)

```

nucleic acid probes for measuring gene expression in a sample derived from human lung comprising single exon nucleic acid probes having one of 12614 nucleic acid sequences mentioned in the specification, or their complements or the 12387 open reading frames derived from the 12614 probes. Also included are a microarray comprising the novel set of probes; the novel set of probes which hybridise at high stringency to a nucleic acid expressed in the human lung; measuring gene expression in a sample derived from human lung, comprising (a) contacting the array with a collection of detectably labeled nucleic acids derived from human lung mRNA, and (b) measuring the label detectably bound to each probe of the array; identifying exons in a eukaryotic genome, comprising (a) algorithmically predicting at least one exon from genomic sequences of the eukaryote; and (b) detecting specific hybridisation of detectably labeled nucleic acids from eukaryote lung mRNA, to a single exon probe, having a fragment identical to the predicted exon, the probe is included in the above mentioned microarray; assigning exons to a single gene, comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberculous scleroidosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension, hyaline membrane disease, open reading frame, ORF.

Alignment Scores:  
Pred. No.: 2,74e-47 Length: 193  
Score: 64.00 Matches: 64  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.48% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABS23881 (1-193)

QY 416 AsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThr 435  
DB 2 AATGTGACCAATGAGTGTGAGATTTTCATTCATTAATGAGTCAGGGTGTTGTGACC 61  
QY 436 ProLysAlaLysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLys 455  
DB 62 CCTAAGGCAAGGAGCAGATGATGAGTGTGGGCACTCTGCACACATTGTACACTGAAG 121  
QY 456 AsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSer 475  
DB 122 AATGGCCGAGTGTGATGCTGTGTCATCTTTGGTCACTGCCCTCTCTATGGATATATAGC 181  
QY 476 AsnValGlnGlu 479  
DB 182 AATGTGCAGGAA 193

RESULT 34  
AAA45505  
ID AAA45505 standard; cDNA; 292 BP.  
XX AC AAA45505;  
XX DT 21-AUG-2000 (first entry)



SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 8.97e-44 Length: 182  
 Score: 60.00 Matches: 60  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.20% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAI24099 (1-182)

QY 849 SerLysLeuThrLeuThrProTTPValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
 DB 182 TCCAAGCTCACCTTAACCCCATGGGTGGCTTCGGAAGATCAATGTCTACTGGTGC 123  
 QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
 DB 122 TGGAGATATGTCCTCCCAATTTACAAATAGTTTACTACAGTGGATCGGCTGAGCCAGT 63  
 QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908  
 DB 62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

RESULT 36

AAI493223/c  
 ID ABA69223 standard; DNA; 182 BP.

XX AC ABA69223;

XX DT 01-FEB-2002 (first entry)

XX DE Human foetal liver single exon nucleic acid probe #17528.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000669.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483447/52.

XX DE Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human foetal liver.

XX PS Claim 4; SEQ ID NO 17528; 639pp + Sequence Listing; English.

XX CC The invention relates to a single exon nucleic acid probe for measuring human gene expression in a sample derived from human foetal liver. The single exon nucleic acid probes may be used for predicting, measuring and displaying gene expression in samples derived from human foetal liver. The present sequence is a single exon nucleic acid probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 8.97e-44 Length: 182  
 Score: 60.00 Matches: 60  
 Percent Similarity: 100.00% Conservativeness: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.20% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABA69223 (1-182)

QY 849 SerLysLeuThrLeuThrProTTPValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
 DB 182 TCCAAGCTCACCTTAACCCCATGGGTGGCTTCGGAAGATCAATGTCTACTGGTGC 123  
 QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
 DB 122 TGGAGATATGTCCTCCCAATTTACAAATAGTTTACTACAGTGGATCGGCTGAGCCAGT 63  
 QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908  
 DB 62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

RESULT 37

AAI49392/c

ID AAI49392 standard; DNA; 182 BP.

XX AC AAI49392;

XX DT 17-OCT-2001 (first entry)

XX DE Probe #18078 used to measure gene expression in human placenta sample.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW Genetic disorder; ss.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488897/53.

XX DE Human genome-derived single exon nucleic acid probes useful for analyzing gene expression in human placenta.

XX PS Claim 25; SEQ ID NO 18078; 654pp; English.

XX CC The present invention relates to single exon nucleic acid probes (SENP). The present sequence is one such probe. The probes are useful for producing a microarray for predicting, measuring and displaying gene expression in samples derived from human placenta. The probes are useful for antenatal diagnosis of human genetic disorders

XX SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:

```

Pred. No.: 8.97e-44 Length: 182
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.20% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAI49392 (1-182)

QY 849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
Db 182 TCCAGCTCACCTTAACCCATGGTGGCGCTTCGGAGATCAATGTGTCTACTGGTGC 123
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuGlnTrpMetProSerGluProSer 888
Db 122 TGGGAAGATATGTCCCATTTACAAATAGTTTACTACAGTGGATCGGCTCAGCCCACT 63
QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
Db 62 GATGCTGGATCTGTGGAATTTATCAGAACCCAGTACTCGGGACTGAAGGCTGCAACC 3

RESULT 38
ABA51213/c
ID ABA51213 standard; DNA; 182 BP.
XX AC
XX ABA51213;
XX AC
DT 01-FEB-2002 (first entry)
XX AC
DE Human breast cell single exon nucleic acid probe #9908.
XX Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer; ss.
XX Homo sapiens.
XX AC
XX WO200157271-A2.
XX AC
XX 09-AUG-2001.
XX AC
XX 30-JAN-2001; 2001WO-US000662.
XX AC
XX 04-FEB-2000; 2000US-0180312P.
XX AC
XX 26-MAY-2000; 2000US-0207456P.
XX AC
XX 30-JUN-2000; 2000US-00608408.
XX AC
XX 03-AUG-2000; 2000US-00632366.
XX AC
XX 21-SEP-2000; 2000US-0234687P.
XX AC
XX 27-SEP-2000; 2000US-0236359P.
XX AC
XX 04-OCT-2000; 2000GB-00024263.
XX AC
(MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-496933/54.
XX DR
XX New spatially-addressable set of single exon nucleic acid probes, useful
XX for measuring gene expression in sample derived from human breast,
XX comprises number of single exon nucleic acid probes.
XX
XX Claim 4; SEQ ID NO 9908; 327pp + Sequence Listing; English.
XX
XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting the
XX probes with a collection of detectably labelled nucleic acids derived
XX from mRNA of human breast, and then measuring the label bound to each
XX probe of the microarray. The probes are useful for verifying the
XX expression of regions of genomic DNA predicted to encode proteins. They
XX are useful for gene discovery, and for determining predisposition and/or
XX prognosing breast disease. Gene expression analysis is useful for
XX assessing the toxicity of chemical agents on cells. The microarray of
XX this invention presents a far greater diversity of probes for measuring

```

```

CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a single exon
CC nucleic acid probe of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;
XX

Alignment Scores:
Pred. No.: 8.97e-44 Length: 182
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.20% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABA51213 (1-182)

QY 849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868
Db 182 TCCAGCTCACCTTAACCCATGGTGGCGCTTCGGAGATCAATGTGTCTACTGGTGC 123
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuGlnTrpMetProSerGluProSer 888
Db 122 TGGGAAGATATGTCCCATTTACAAATAGTTTACTACAGTGGATCGGCTCAGCCCACT 63
QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
Db 62 GATGCTGGATCTGTGGAATTTATCAGAACCCAGTACTCGGGACTGAAGGCTGCAACC 3

RESULT 39
ABA36140/c
ID ABA36140 standard; DNA; 182 BP.
XX AC
XX ABA36140;
XX AC
DT 23-JAN-2002 (first entry)
XX AC
XX Probe #14606 for gene expression analysis in human heart cell sample.
XX Human; gene expression; heart; microarray; vascular system; probe;
XX cardiovascular disease; hypertension; cardiac arrhythmia;
XX congenital heart disease; ss.
XX Homo sapiens.
XX WO200157274-A2.
XX AC
XX 09-AUG-2001.
XX AC
XX 30-JAN-2001; 2001WO-US000666.
XX AC
XX 04-FEB-2000; 2000US-0180312P.
XX AC
XX 26-MAY-2000; 2000US-0207456P.
XX AC
XX 30-JUN-2000; 2000US-00608408.
XX AC
XX 03-AUG-2000; 2000US-00632366.
XX AC
XX 21-SEP-2000; 2000US-0234687P.
XX AC
XX 27-SEP-2000; 2000US-0236359P.
XX AC
XX 04-OCT-2000; 2000GB-00024263.
XX AC
(MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX DR
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX
XX Claim 4; SEQ ID NO 14606; 530pp; English.
XX

```

CC The present invention relates to single exon nucleic acid probes for  
 CC measuring human gene expression in a sample derived from human heart. The  
 CC present sequence is one such probe. The probes may be used for  
 CC predicting, measuring and displaying gene expression in samples derived  
 CC from the human heart via microarrays. By measuring gene expression, the  
 CC probes are useful for predicting, diagnosing, grading, staging,  
 CC monitoring and prognosing diseases of the human heart and vascular system  
 CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and  
 CC congenital heart disease. Note: The sequence data for this patent did not  
 CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 8.97e-44 Length: 182  
 Score: 60.00 Matches: 60  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.20% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABA36140 (1-182)

QY 849 SerLysLeuThrLeuThrProTTPValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
 DB 182 TCCAAGCTCACCTTAACCCCATGGGTCGGCTTCGGAAGATCAATGTCTACTGGTGC 123  
 QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
 DB 122 TGGGAAGATATGCCCCCATTTACAAATAGTTTACTACAGTGGATCGGCTGAGCCAGT 63  
 QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaLaThr 908  
 DB 62 GATGCTGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAGGCTGCAACC 3

RESULT 40  
 AAK43324/c  
 ID AAK43324 standard; DNA; 182 BP.  
 XX  
 AC AAK43324;  
 XX  
 DT 06-NOV-2001 (first entry)  
 XX  
 DE Human bone marrow expressed single exon probe SEQ ID NO: 17881.  
 XX  
 KW Human; bone marrow expressed exon; gene expression analysis; probe;  
 KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200157276-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000668.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-488900/53.  
 XX  
 PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human bone marrow.

XX Example 4; SEQ ID NO 17881; 658pp + Sequence Listing; English.  
 PS  
 CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC bone marrow. They can be used to measure gene expression in bone marrow  
 CC samples, which may enable the improved diagnosis and treatment of cancers  
 CC such as lymphoma, leukaemia and myeloma. The present sequence is one of  
 CC the probes of the invention  
 XX  
 SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 8.97e-44 Length: 182  
 Score: 60.00 Matches: 60  
 Percent Similarity: 100.00% Conservatives: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.20% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAK43324 (1-182)

QY 849 SerLysLeuThrLeuThrProTTPValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
 DB 182 TCCAAGCTCACCTTAACCCCATGGGTCGGCTTCGGAAGATCAATGTCTACTGGTGC 123  
 QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
 DB 122 TGGGAAGATATGCCCCCATTTACAAATAGTTTACTACAGTGGATCGGCTGAGCCAGT 63  
 QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaLaThr 908  
 DB 62 GATGCTGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAGGCTGCAACC 3

RESULT 41  
 AAK17517/c  
 ID AAK17517 standard; DNA; 182 BP.  
 XX  
 AC AAK17517;  
 XX  
 DT 05-NOV-2001 (first entry)  
 XX  
 DE Human brain expressed single exon probe SEQ ID NO: 17508.  
 XX  
 KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
 KW ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO200157275-A2.  
 XX  
 PD 09-AUG-2001.  
 XX  
 PF 30-JAN-2001; 2001WO-US000667.  
 XX  
 PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-00608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX

PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn SG, Hanzel DK, Chen W, Rank DR;  
 XX  
 DR WPI; 2001-483446/52.  
 XX  
 PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains.  
 XX



PS Example 4; SEQ ID NO 17508; 650pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid  
XX probes which are derived from genomic sequences expressed in the human  
CC brain. They can be used to measure gene expression in brain cell samples,  
CC which may enable the diagnosis and improved treatment of nervous system  
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
CC epilepsy and cancers. The present sequence is one of the probes of the  
CC invention

XX SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 8,978-44 Length: 182  
Score: 60.00 Matches: 60  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.20% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAK17517 (1-182)

QY 849 SerLysLeuThrLeuThrProTrrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
DB 182 TCCAGCTCACCTTAACCCCATGGTGGCCTTCGGAGATCAATGTGTCTACTGTGTC 123  
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
DB 122 TGGAGAGATATGTCCTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCAGT 63  
QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908  
DB 62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAGGCTGCAACC 3

#### RESULT 42

AAI09674/c

ID ABS42949 standard; DNA; 182 BP.

XX AC ABS42949;

XX 25-FEB-2003 (first entry)

XX Human liver single exon probe, SEQ ID No 17939.

XX Human: single exon nucleic acid probe; liver; cirrhosis;  
KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;  
KW coronary heart disease; ss.

XX Homo sapiens.

XX WO200157273-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000664.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing  
PT gene expression in human adult liver.

PS Claim 4; SEQ ID NO 17939; 658pp; English.

XX The invention relates to a single exon nucleic acid probe (SENP) (I) for  
XX measuring human gene expression in a sample derived from human adult  
CC liver, comprising one of 13109 defined nucleotide sequences given in the  
CC specification (or complements/ fragments). The probe hybridises at high  
CC stringency to a nucleic acid molecule expressed in the human adult liver.  
CC (I) may be used for predicting, measuring and displaying gene expression  
CC in samples derived from human adult liver. The genes identified may be  
CC involved in genetic liver diseases such as cirrhosis,  
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
CC associated with coronary heart disease. ABS25011-ABS51005 represent human  
CC liver single exon nucleic acid probes of the invention. Note: The  
CC sequence information for this patent does not appear in the printed  
CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 8,978-44 Length: 182  
Score: 60.00 Matches: 60  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.20% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABS42949 (1-182)

QY 849 SerLysLeuThrLeuThrProTrrpValGlyLeuArgLysIleAsnValSerTyrTrpCys 868  
DB 182 TCCAGCTCACCTTAACCCCATGGTGGCCTTCGGAGATCAATGTGTCTACTGTGTC 123  
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888  
DB 122 TGGAGAGATATGTCCTTACAAATAGTTTACTACAGTGGATGCCGTCTGAGCCAGT 63  
QY 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908  
DB 62 GATGCTGGATTCTGTGGAATTTTATCAGAACCCAGTACTCGGGGACTGAGGCTGCAACC 3

#### RESULT 43

AAI09674/c

ID AAI09674 standard; DNA; 182 BP.

XX AC AAI09674;

XX 09-OCT-2001 (first entry)

XX Probe #9665 used to measure gene expression in human breast sample.

XX Probe; human; breast disease; breast cancer; development disorder; ss;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX Homo sapiens.

XX WO200157270-A2.

XX 09-AUG-2001.

XX 29-JAN-2001; 2001WO-US000661.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.  
 XX Novel single exon nucleic acid probe used to measuring gene expression in  
 PT a human breast.  
 XX Claim 25; SEQ ID NO 9665; 322pp; English.  
 XX The present invention relates to novel single exon nucleic acid probes.  
 CC The present sequence is one such probe. The probes are useful for  
 CC measuring human gene expression in a human breast sample, where the probe  
 CC hybridises at high stringency to a nucleic acid expressed in the human  
 CC breast. The probes are useful for predicting, diagnosing, grading,  
 CC staging, monitoring and prognosing diseases of the human breast,  
 CC particularly those diseases with polygenic aetiology. The diseases  
 CC include: breast cancer, disorders of development, inflammatory diseases  
 CC of the breast, fibrocystic changes, proliferative breast disease and non-  
 CC carcinoma tumours. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 8,97e-44 Length: 182  
 Score: 60.00 Matches: 60  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.20% Indels: 0  
 DB: 5 Gaps: 0  
 US-09-787-097-12 (1-1429) x AAI09674 (1-182)  
 Qy 849 SerLysLeuThrLeuThrProTTPValGlyLeuArgLysLeuValSerTyrTrpCys 868  
 Db 182 TCCAGCTCACCTTAACCCATGCGGCTTCGGAAGATCAATGTGCTACTGGTGC 123  
 Qy 869 TrpGluAspMetSerProPheThrAnSerLeuLeuGlnTrpMetProSerGluProSer 888  
 Db 122 TGGAGAGATATGCCCATTTACAAATAGTTTACTACAGTGGATCGGCTCAGGCCAGT 63  
 Qy 889 AspAlaGlyPheCysGlyLleLeuSerGluProSerThrArgGlyLeuLysAlaThr 908  
 Db 62 GATGCTGGATCTGTGGATTTTATCAGAACCCAGTACTCGGGACTGAGGCTGCACC 3  
 RESULT 44  
 ABS17404/c  
 ID ABS17404 standard; DNA; 182 BP.  
 XX AC ABS17404;  
 XX DT 19-AUG-2002 (first entry)  
 XX Human genome-derived single exon probe ORF from lung SEQ ID No 17395.  
 DE Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberculous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease; open reading frame; ORF.  
 XX Homo sapiens.  
 OS WO200186003-A2.  
 FN 15-NOV-2001.  
 XX 30-JAN-2001; 2001WO-US0000665.

PR 04-FEB-2000; 2000US-0180312P.  
 PR 26-MAY-2000; 2000US-0207456P.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-00632366.  
 PR 21-SEP-2000; 2000US-0234687P.  
 PR 27-SEP-2000; 2000US-0236359P.  
 PR 04-OCT-2000; 2000GB-00024263.  
 XX (MOLE-) MOLECULAR DYNAMICS INC.  
 PA Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2002-114183/15.  
 XX Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.  
 XX Claim 4; SEQ ID NO 17395; 634pp; English.  
 CC The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridise at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberculous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a single exon probe open reading frame of the  
 CC invention. Note: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format directly  
 CC from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX SQ Sequence 182 BP; 48 A; 46 C; 45 G; 43 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 8,97e-44 Length: 182  
 Score: 60.00 Matches: 60  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 4.20% Indels: 0  
 DB: 6 Gaps: 0  
 US-09-787-097-12 (1-1429) x ABS17404 (1-182)  
 Qy 849 SerLysLeuThrLeuThrProTTPValGlyLeuArgLysLeuValSerTyrTrpCys 868  
 Db 182 TCCAGCTCACCTTAACCCATGCGGCTTCGGAAGATCAATGTGCTACTGGTGC 123

QY 869 TtpGluAspMetSerProPhetHrAsnSerLeuLeuGlnTtpMetProSerGluProSer 888  
 Db 122 TGGAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGCTGTGAGCCAGT 63  
 QY 889 AspAlaGlyPheCysGlyLeuSerGluProSerThrArgGlyLeuLeuAlaAlaThr 908  
 Db 62 GATGCTGGATTCGTGGATTTATCAGAACCCAGTACTCGGGGACTGAAGGCTGCAACC 3

## RESULT 45

AA114910/C  
 ID AA114910 standard; DNA; 375 BP.

XX AC AA114910;

XX 12-OCT-2001 (first entry)

XX DE Probe #4843 for gene expression analysis in human cervical cell sample.

XX KW Probe; human; microarray; gene expression; cervical epithelial cell;

XX KW cervical cancer; ss.

XX OS Homo sapiens.

XX PN WO200157278-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000670.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-488901/53.

XX DR Human genome-derived single exon nucleic acid probes useful for analyzing

XX PT gene expression in human cervical epithelial cells.

XX PS Claim 25; SEQ ID NO 4843; 487bp; English.

XX CC The present invention relates to human single exon nucleic acid probes  
 CC (SENP). The present sequence is one such probe. The SENPs are derived  
 CC from human HeLa cells. The SENPs can be used to produce a single exon  
 CC sample derived from human cervical epithelial cells. By measuring gene  
 CC expression, the probes are therefore useful in grading and/or staging of  
 CC diseases of the cervix, notably cervical cancer. Note: The sequence data  
 CC for this patent did not form part of the printed specification, but was  
 CC obtained in electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 7.92e-41 Length: 375  
 Score: 57.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.99% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AA114910 (1-375)

QY 865 SerTyrTpcCystTpcGluAspMetSerProPhetHrAsnSerLeuLeuGlnTtpMetPro 884

Db 374 TCCTACTGGTGGGAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATGCGC 315  
 QY 885 SerGluProSerAspAlaGlyPheCysGlyLeuSerGluProSerThrArgGlyLeu 904  
 Db 314 TCTGAGCCAGTGGATGCTGGATTTCTGTGGATTTATCAGAACCCAGTACTCGGGACTG 255  
 QY 905 LysAlaAlaThrCysLeuAsnProLeuAsnGlySerValCysGluArgPro 921  
 Db 254 AAGGCTGCAACCTGCATCAACCCACTCAATGCTAGTGTCTGTGAAAGGCCT 204

## RESULT 46

ABA56648/c

ID ABA56648 standard; DNA; 375 BP.

XX AC ABA56648;

XX 01-FEB-2002 (first entry)

XX DE Human foetal liver single exon nucleic acid probe #4953.

XX KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX OS Homo sapiens.

XX PN WO200157277-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000669.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX PA Penn SG, Hanzel DK, Chen W, Rank DR;

XX PI WPI; 2001-483447/52.

XX DR Human genome-derived single exon nucleic acid probes useful for analyzing

XX PT gene expression in human fetal liver.

XX PS Claim 1; SEQ ID NO 4953; 639pp + Sequence Listing; English.

XX CC The invention relates to a single exon nucleic acid probe for measuring  
 CC human gene expression in a sample derived from human foetal liver. The  
 CC single exon nucleic acid probes may be used for predicting, measuring and  
 CC displaying gene expression in samples derived from human fetal liver. The  
 CC present sequence is a single exon nucleic acid probe of the invention.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 7.92e-41 Length: 375  
 Score: 57.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.99% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABA56648 (1-375)

QY 865 SerTyrTpcCystTpcGluAspMetSerProPhetHrAsnSerLeuLeuGlnTtpMetPro 884

```

Db 374 TCCTACTGGTGGGAGATATGTCCCATTTTACAAATAGTTTACTACAGTGGATGCG 315
QY 885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904
Db 314 TCTGAGCCAGTGTGATGCTGGATTTCTGTGGATTTTATCAGAACCCAGTACTCGGGAGCTG 255
QY 905 LysAlaAlaThrCysIleLeuProLeuAsnGlySerValCysGluArgPro 921
Db 254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCT 204

RESULT 47
AAI36262/c
ID AAI36262 standard; DNA; 375 BP.
XX
AC AAI36262;
XX
DT 17-OCT-2001 (first entry)
XX
DE Probe #4948 used to measure gene expression in human placenta sample.
XX
KW Probe; microarray; human; placenta; antenatal diagnosis;
XX genetic disorder; ss.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000663.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-488897/53.
XX
DR Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human placenta.
XX
PS Claim 25; SEQ ID NO 4948; 654bp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP).
XX The present sequence is one such probe. The probes are useful for
XX producing a microarray for predicting, measuring and displaying gene
XX expression in samples derived from human placenta. The probes are useful
XX for antenatal diagnosis of human genetic disorders
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7,92e-41 Length: 375
Score: 57.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.99% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAI36262 (1-375)

QY 865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuGlnTrpMetPro 884
Db 374 TCCTACTGGTGGGAGATATGTCCCATTTTACAAATAGTTTACTACAGTGGATGCG 315
QY 885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904

```

```

Db 314 TCTGAGCCAGTGTGATGCTGGATTTCTGTGGAATTTATCAGAACCCAGTACTCGGGAGCTG 255
QY 905 LysAlaAlaThrCysIleLeuProLeuAsnGlySerValCysGluArgPro 921
Db 254 AAGGCTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCT 204

RESULT 48
ABA46100/c
ID ABA46100 standard; DNA; 375 BP.
XX
AC ABA46100;
XX
DT 01-FEB-2002 (first entry)
XX
DE Human breast cell single exon nucleic acid probe #4795.
XX
KW Human; microarray; single exon probe; gene expression; breast; disease;
XX cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157271-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000662.
XX
PR 04-FEB-2000; 2000US-0180312P.
XX
PR 26-MAY-2000; 2000US-0207456P.
XX
PR 30-JUN-2000; 2000US-00608408.
XX
PR 03-AUG-2000; 2000US-00632366.
XX
PR 21-SEP-2000; 2000US-0234687P.
XX
PR 27-SEP-2000; 2000US-0236359P.
XX
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
WPI; 2001-496933/54.
XX
DR New spatially-addressable set of single exon nucleic acid probes, useful
XX for measuring gene expression in sample derived from human breast,
XX comprises number of single exon nucleic acid probes.
XX
PS Claim 1; SEQ ID NO 4795; 327pp + Sequence Listing; English.
XX
CC The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human breast and BT 474 cells. The method involves contacting the
XX probes with a collection of detectably labelled nucleic acids derived
XX from mRNA of human breast, and then measuring the label bound to each
XX probe of the microarray. The probes are useful for verifying the
XX expression of regions of genomic DNA predicted to encode proteins. They
XX are useful for gene discovery, and for determining predisposition and/or
XX assessing the toxicity of chemical agents on cells. The microarray of
XX this invention presents a far greater diversity of probes for measuring
XX gene expression, with far less bias than expressed sequence tag
XX microarrays. The method is suitable for rapid production of functional
XX information from genomic sequence. The present sequence is a single exon
XX nucleic acid probe of the invention. Note: The sequence data for this
XX patent did not form part of the printed specification, but was obtained
XX in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7,92e-41 Length: 375
Score: 57.00 Matches: 57
Percent Similarity: 100.00% Conservative: 0

```

```
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.99% Indels: 0
DB: 4 Gaps: 0
US-09-787-097-12 (1-1429) x ABA46100 (1-375)
QY 865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuGlnTrpMetPro 884
DB 374 TCCTACTGGTCTGGGAAGATATGTCCTCCATTACAAATAGTTTACTACAGTGGATGCCG 315
QY 885 SerGluProSerAspAlaGlyPheCysGlyLeuSerGluProSerThrArgGlyLeu 904
DB 314 TCTGAGCCCGATGCTGGATTCTGTGGAATTTATCAGACCCAGTACTCGGGGACTG 255
QY 905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
DB 254 AAGGCTGCACTGCTGCACTCAACCCACTCAATGGTAGTGTCTGTGAAGGCCT 204
RESULT 49
ABA26254/c
ID ABA26254 standard; DNA; 375 BP.
XX AC ABA26254;
XX DT 23-JAN-2002 (first entry)
XX DE Probe #4720 for gene expression analysis in human heart cell sample.
XX KW Human; gene expression; heart; microarray; vascular system; probe;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease; ss.
XX OS Homo sapiens.
XX PN WC200157274-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000666.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-48899/53.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX hearts.
XX PS Claim 1; SEQ ID NO 4720; 530pp; English.
XX CC The present invention relates to single exon nucleic acid probes for
XX CC measuring human gene expression in a sample derived from human heart. The
XX CC present sequence is one such probe. The probes may be used for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from the human heart via microarrays. By measuring gene expression, the
XX CC probes are useful for predicting, diagnosing, grading, staging,
XX CC monitoring and prognosing diseases of the human heart and vascular system
XX CC e.g. cardiovascular disease, hypertension, cardiac arrhythmias and
XX CC congenital heart disease. Note: The sequence data for this patent did not
XX CC form part of the printed specification, but was obtained in electronic
XX CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;
Alignment Scores: 7.92e-41 Length: 375
Pred. No.: 375
```

Score: 57.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.99% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAK30307 (1-375)

QY 865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884  
 DB 374 TCTACTGGTGGCGGGAAGATATGCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315  
 QY 885 SerGluProSerAspAlaGlyPheCysGlyLeuLeuSerGluProSerThrArgGlyLeu 904  
 DB 314 TCTGAGCCAGTGTGCTGGAATTTCTGGAATTTATCAGAACCCAGTACTCGGGACTG 255  
 QY 905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
 DB 254 AAGGCTGCAACCTGCATCAACCACTCAATGGTAGTGTCTGTGAAGGCCT 204

## RESULT 51

AAK04785/c  
 ID AAK04785 standard; DNA; 375 BP.

XX AC AAK04785;

XX DT 05-NOV-2001 (first entry)

XX DE Human brain expressed single exon probe SEQ ID NO: 4776.

XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;  
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;  
 KW ss.

XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000667.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOL- ) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human  
 PT brains.

XX PS Example 4; SEQ ID NO 4776; 650pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid  
 CC probes which are derived from genomic sequences expressed in the human  
 CC brain. They can be used to measure gene expression in brain cell samples,  
 CC which may enable the diagnosis and improved treatment of nervous system  
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,  
 CC epilepsy and cancers. The present sequence is one of the probes of the  
 CC invention

XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.92e-41 Length: 375

Score: 57.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.99% Indels: 0  
 DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAK04785 (1-375)

QY 865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884  
 DB 374 TCTACTGGTGGCGGGAAGATATGCCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315  
 QY 885 SerGluProSerAspAlaGlyPheCysGlyLeuLeuSerGluProSerThrArgGlyLeu 904  
 DB 314 TCTGAGCCAGTGTGCTGGAATTTCTGGAATTTATCAGAACCCAGTACTCGGGACTG 255  
 QY 905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
 DB 254 AAGGCTGCAACCTGCATCAACCACTCAATGGTAGTGTCTGTGAAGGCCT 204

## RESULT 52

ABS29950/c  
 ID ABS29950 standard; DNA; 375 BP.

XX AC ABS29950;

XX DT 25-FEB-2003 (first entry)

XX DE Human liver single exon probe, SEQ ID No 4940.

XX KW Human; single exon nucleic acid probe; liver; cirrhosis;  
 KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;  
 KW coronary heart disease; ss.

XX OS Homo sapiens.

XX PN WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000664.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOL- ) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488898/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing  
 PT gene expression in human adult liver.

XX PS Claim 1; SEQ ID NO 4940; 658pp; English.

XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for  
 CC measuring human gene expression in a sample derived from human adult  
 CC liver, comprising one of 13109 defined nucleotide sequences given in the  
 CC specification (or complements/ fragments). The probe hybridizes at high  
 CC stringency to a nucleic acid molecule expressed in the human adult liver.  
 CC (I) may be used for predicting, measuring and displaying gene expression  
 CC in samples derived from human adult liver. The genes identified may be  
 CC involved in genetic liver diseases such as cirrhosis,  
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is  
 CC associated with coronary heart disease. ABS25011-ABS51005 represent human  
 CC liver single exon nucleic acid probes of the invention. Note: The  
 CC sequence information for this patent does not appear in the printed

CC specification but was obtained in electronic format directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences  
XX  
SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:  
Pred. No.: 7.92e-41 Length: 375  
Score: 57.00 Matches: 57  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.99% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABS29950 (1-375)

QY 865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884  
DB 374 TCCTACTGTTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315  
QY 885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904  
DB 314 TCTGAGCCAGTGGATGCTGGATTTCTGTGGAATTTATCAGAACCCAGTACTCGGGGACTG 255  
QY 905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
DB 254 AAGGCTGCAACCTGCATCAACCCACTCAATGTTGTCTGTGAAAGGCCT 204

RESULT 53

AA104689/c

ID AA104689 standard; DNA; 375 BP.

XX

AC AA104689;

XX

DT 09-OCT-2001 (first entry)

XX

DE Probe #4680 used to measure gene expression in human breast sample.

XX

KW Probe; human; breast disease; breast cancer; development disorder; ss;

KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.

XX

OS Homo sapiens.

XX

FN WO200157270-A2.

XX

PD 09-AUG-2001.

XX

PF 29-JAN-2001; 2001WO-US000661.

XX

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.

XX

PT Novel single exon nucleic acid probe used to measuring gene expression in  
PT a human breast.

XX

PS Claim 25; SEQ ID NO 4680; 322pp; English.

XX

CC The present invention relates to novel single exon nucleic acid probes.

CC The present sequence is one such probe. The probes are useful for

CC measuring human gene expression in a human breast sample, where the probe

CC hybridises at high stringency to a nucleic acid expressed in the human

CC breast. The probes are useful for predicting, diagnosing, grading,

CC staging, monitoring and prognosing diseases of the human breast,

CC

CC

CC

CC

CC

CC

CC

CC particularly those diseases with polygenic aetiology. The diseases  
CC include: breast cancer, disorders of development, inflammatory diseases  
CC of the breast, fibrocystic changes, proliferative breast disease and non-  
CC carcinoma tumours. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
XX

SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.92e-41 Length: 375

Score: 57.00 Matches: 57

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.99% Indels: 0

DB: 5 Gaps: 0

US-09-787-097-12 (1-1429) x AA104689 (1-375)

QY 865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884

DB 374 TCCTACTGTTGCTGGGAAGATATGTCCTCCCATTTACAAATAGTTTACTACAGTGGATGCCG 315

QY 885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904

DB 314 TCTGAGCCAGTGGATGCTGGATTTCTGTGGAATTTATCAGAACCCAGTACTCGGGGACTG 255

QY 905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921

DB 254 AAGGCTGCAACCTGCATCAACCCACTCAATGTTGTCTGTGAAAGGCCT 204

RESULT 54

ABS04892/c

ID ABS04892 standard; DNA; 375 BP.

XX

AC ABS04892;

XX

DT 19-AUG-2002 (first entry)

XX

DE Human genome-derived single exon probe from lung SEQ ID No 4883.

XX

KW Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;

KW chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagenier syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX

OS Homo sapiens.

XX

PN WO200186003-A2.

XX

PD 15-NOV-2001.

XX

XX 30-JAN-2001; 2001WO-US000665.

XX

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PA

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



PT Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.

XX Claim 1; SEQ ID NO 4883; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridise at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA, and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying exons from genomic sequence by the method  
 CC above and (b) measuring the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagenen syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a single exon probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 375 BP; 101 A; 95 C; 63 G; 116 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 7 92e-41 Length: 375  
 Score: 57.00 Matches: 57  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.99% Indels: 0  
 DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x ABS04892 (1-375)

QY 865 SerTyTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884  
 Db 374 TCCTACTGGTGTGGGAAGATATGTCCTCCATTTACAAATAGTTTACTACAGTGGATCGG 315  
 QY 885 SerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904  
 Db 314 TCTGACCCAGTGTGCTGGATCTCTGGAAATTTATCAGACCCAGTACTCGGGGACTG 255

RESULT 55

AZ291926/c

ID AA291926 standard; DNA; 226 BP.

XX

AC AA291926;

XX 08-JUN-2000 (first entry)  
 XX Wild type murine mahogany gene exon 15 sequence.  
 XX Mahogany gene; mouse; mg gene; regulatory defect; gene therapy; obesity;  
 KW weight regulation; cell therapy; body weight disorder; cachexia;  
 KW anorexia; hyperpigmentation; increased metabolic rate disorder;  
 KW hyperphagia; Antioesity; antianorexic; anticachexic; ds.  
 OS Mus sp.  
 XX WO200005373-A2.  
 XX 03-FEB-2000.  
 XX 21-JUL-1999; 99WO-US016484.  
 XX 21-JUL-1998; 98US-0093630P.  
 PR 20-OCT-1998; 98US-0104978P.  
 PR 05-FEB-1999; 99US-00245041.  
 XX (MILL-) MILLENIUM PHARM INC.  
 XX Moore K, Nagle DL;  
 XX WPI; 2000-195103/17.  
 XX New human and murine mahogany genes, useful, e.g. for diagnosis and  
 PT treatment of body weight disorders.  
 XX Example; Fig 6; 188pp; English.

XX This sequence is a fragment of a murine mahogany gene of the invention.  
 CC The mahogany genes are used: (i) to produce recombinant mahogany (mg)  
 CC proteins (II); (ii) as a source of antisense, ribozyme or triplex-forming  
 CC therapeutics; (iii) as a source of diagnostic probes and primers for  
 CC detecting expression of mg genes or mutations, regulatory defects, in  
 CC this gene, or for isolation of related sequences; and (iv) in (cell-  
 CC based) gene therapy. (II) are used to raise specific antibodies (Ab); to  
 CC identify other (extra)cellular products involved in weight regulation,  
 CC and to screen for agents that disrupt interaction between (II) and other  
 CC macromolecules. The Ab are used to detect abnormal levels (or function)  
 CC of (II) (for diagnosis, prognosis or monitoring of treatment); to  
 CC evaluate (II)-expressing cells intended for cell therapy, and as  
 CC therapeutic mg inhibitors. Cells that express the mg gene (or contain the  
 CC mg polypeptide) are used to identify agents (A) that modulate mg  
 CC activity (A) are potentially useful for the treatment of body weight  
 CC disorders, particularly obesity, cachexia or anorexia, or other  
 CC conditions associated with the mg gene such as hyperpigmentation,  
 CC hyperphagia and disorders that result in increased metabolic rate  
 XX SQ Sequence 226 BP; 51 A; 66 C; 52 G; 57 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 1 01e-35 Length: 226  
 Score: 51.00 Matches: 51  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.57% Indels: 0  
 DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x AA291926 (1-226)

QY 871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890  
 Db 225 GATATGTCCTCAATTCACAAATAGTTTCTGTCAGTGGATGTCATCTGAGCCCATGTGCT 166  
 QY 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaThrCysIle 910  
 Db 165 GGCTTCTGTGGATCTTGTTCAGAGCTTAGTACTCGGGGATTAAGGCTCAACTGCATC 106  
 QY 911 AsnProLeuAsnGlySerValCysGluArgPro 921

```
Db 105 AACCTTCTCATGGCAGGCTCTGTGAAAGGCT 73
|||||
RESULT 56
AAS69062
ID AAS69062 standard; cDNA; 2267 BP.
AC AAS69062;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #4866.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX
FN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
XX WPI; 2001-639362/73.
DR P-PSDB; ABG04875.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 1; SEQ ID NO 4866; 103pp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
CC coding sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 2267 BP; 917 A; 455 C; 453 G; 442 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 5.04e-33 Length: 2267
Score: 49.00 Matches: 49
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.43% Indels: 0
DB: 5 Gaps: 0

US-09-787-097-12 (1-1429) x AAS69062 (1-2267)
Qy 1140 CysGluValGluAsnArgTyrGlnGlyAsnProLeuArgGlyThrCysTyrThrLeu 1159
Db 163 TGTGAGGTAGAAATCGTATACCAAGGAAACCTCTCAGAGGACATGTTATTACTCTT 222
Qy 1160 LeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrThrAla 1179
Db 223 CTTATTGACTATCAGTTCACTTTAGTCTATCCCGAGAGATGATCGCTATTACAGCT 282
Qy 1180 IleAsnPheValAlaThrProAspGlu 1188
Db 283 ATCAATTTTGTGGCTACTCTCTGACGAA 309
RESULT 57
AAI13789/c
ID AAI13789 standard; DNA; 402 BP.
AC AAI13789;
XX
DT 12-OCT-2001 (first entry)
XX
DE Probe #3722 for gene expression analysis in human cervical cell sample.
XX
KW Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer; ss.
XX
OS Homo sapiens.
XX
PN WO200157278-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US000670.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human cervical epithelial cells.
XX
PS Claim 25; SEQ ID NO 3722; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENPs). The present sequence is one such probe. The SENPs are derived
CC from human HeLa cells. The SENPs can be used to produce a single exon
CC microarray, which can be used for measuring human gene expression in a
CC sample derived from human cervical epithelial cells. By measuring gene
CC expression, the probes are therefore useful in grading and/or staging of
CC diseases of the cervix, notably cervical cancer. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4.58e-31 Length: 402
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.22% Indels: 0
```

DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAI13789 (1-402)

QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081

Db 389 GCTTGCCTAATGCAACGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrglyAspProThr 1101

Db 329 AACCTGACACAGCAGCAGCACTGCGAGACCTGCTGATATCTGGCTTCTACGGTGTATCCACC 270

QY 1102 AsnGlyGlyLysCysGln 1107

Db 269 AATGGAGGGAATGTCTAG 252

RESULT 59

AAI35149/C

ID AAI35149 standard; DNA; 402 BP.

XX

AC AAI35149;

DT 01-FEB-2002 (first entry)

XX

DE Human foetal liver single exon nucleic acid probe #3804.

XX

KW Human; foetal liver; gene expression; single exon nucleic acid probe; ss.

XX

OS Homo sapiens.

XX

PN WO200157277-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US0000669.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-483447/52.

XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human foetal liver.

XX

PS Claim 1; SEQ ID NO 3804; 639pp + Sequence Listing; English.

XX

CC The invention relates to a single exon nucleic acid probe for measuring

CC human gene expression in a sample derived from human foetal liver. The

CC single exon nucleic acid probes may be used for predicting, measuring and

CC displaying gene expression in samples derived from human foetal liver. The

CC present sequence is a single exon nucleic acid probe of the invention.

CC Note: The sequence data for this patent did not form part of the printed

CC specification, but was obtained in electronic format directly from WIPO

CC at ftp.wipo.int/pub/published\_pct\_sequences

XX

SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.58e-31 Length: 402

Score: 46.00 Matches: 46

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.22% Indels: 0

DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABA55499 (1-402)

QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081

Db 389 GCTTGCCTAATGCAACGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrglyAspProThr 1101

Db 329 AACCTGACACAGCAGCAGCACTGCGAGACCTGCTGATATCTGGCTTCTACGGTGTATCCACC 270

QY 1102 AsnGlyGlyLysCysGln 1107

Db 269 AATGGAGGGAATGTCTAG 252

RESULT 59

AAI35149/C

ID AAI35149 standard; DNA; 402 BP.

XX

AC AAI35149;

DT 17-OCT-2001 (first entry)

XX

DE Probe #3835 used to measure gene expression in human placenta sample.

XX

KW Probe; microarray; human; placenta; antenatal diagnosis;

XX

OS Homo sapiens.

XX

PN WO200157272-A2.

XX

PD 09-AUG-2001.

XX

PF 30-JAN-2001; 2001WO-US0000663.

XX

PR 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.

PR 27-SEP-2000; 2000US-0236359P.

PR 04-OCT-2000; 2000GB-00024263.

XX

PA (MOLE-) MOLECULAR DYNAMICS INC.

XX

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX

DR WPI; 2001-488897/53.

XX

PT Human genome-derived single exon nucleic acid probes useful for analyzing

PT gene expression in human placenta.

XX

PS Claim 25; SEQ ID NO 3835; 654pp; English.

XX

CC The present invention relates to single exon nucleic acid probes (SENP).

CC The present sequence is one such probe. The probes are useful for

CC producing a microarray for predicting, measuring and displaying gene

CC expression in samples derived from human placenta. The probes are useful

CC for antenatal diagnosis of human genetic disorders

XX

SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4.58e-31 Length: 402

Score: 46.00 Matches: 46

Percent Similarity: 100.00% Conservatives: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.22% Indels: 0

DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAI35149 (1-402)



CC form part of the printed specification, but was obtained in electronic  
 CC format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4.58e-31 Length: 402  
 Score: 46.00 Matches: 46  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.22% Indels: 0  
 DB: Gaps: 0

US-09-787-097-12 (1-1429) x ABA25215 (1-402)

QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081

Db 389 GCTTGCCAAATGCAACGGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrglyAspProThr 1101

Db 329 AACCTGACCACAGCAAGCACTGCGAGACTGCATATCTGGCTTCTACGGTGATCCACC 270

QY 1102 AsnGlyGlyLysCysGln 1107

Db 269 AATGGAGGGAATGTCAG 252

RESULT 62

AAK29197/c

ID AAK29197 standard; DNA; 402 BP.

XX AC AAK29197;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed single exon probe SEQ ID NO: 3754.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;

XX KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000668.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX PT gene expression in human bone marrow.

XX PS Example 4; SEQ ID NO 3754; 659pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC bone marrow. They can be used to measure gene expression in bone marrow

XX CC samples, which may enable the improved diagnosis and treatment of cancers

XX CC such as lymphoma, leukaemia and myeloma. The present sequence is one of

XX CC the probes of the invention

XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4.58e-31 Length: 402  
 Score: 46.00 Matches: 46  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.22% Indels: 0  
 DB: Gaps: 0

US-09-787-097-12 (1-1429) x AAK29197 (1-402)

QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081

Db 389 GCTTGCCAAATGCAACGGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330

QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrglyAspProThr 1101

Db 329 AACCTGACCACAGCAAGCACTGCGAGACTGCATATCTGGCTTCTACGGTGATCCACC 270

QY 1102 AsnGlyGlyLysCysGln 1107

Db 269 AATGGAGGGAATGTCAG 252

RESULT 63

AAK03733/c

ID AAK03733 standard; DNA; 402 BP.

XX AC AAK03733;

XX DT 05-NOV-2001 (first entry)

XX DE Human brain expressed single exon probe SEQ ID NO: 3724.

XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;

XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;

XX OS Homo sapiens.

XX PN WO200157275-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000667.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-483446/52.

XX PT Single exon nucleic acid probes for analyzing gene expression in human

XX PT brains.

XX PS Example 4; SEQ ID NO 3724; 650pp + Sequence Listing; English.

XX CC The present invention provides a number of single exon nucleic acid

XX CC probes which are derived from genomic sequences expressed in the human

XX CC brain. They can be used to measure gene expression in brain cell samples,

XX CC which may enable the diagnosis and improved treatment of nervous system

XX CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

XX CC epilepsy and cancers. The present sequence is one of the probes of the

XX CC invention

```
XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4,58e-31 Length: 402
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.22% Indels: 0
DB: 4 Gaps: 0
US-09-787-097-12 (1-1429) x AAK03733 (1-402)
QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
DB 389 GCTTGCCAATGCAACGGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330
QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrlGlyAspProThr 1101
DB 329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCACC 270
QY 1102 AsnGlyGlyLysCysGln 1107
DB 269 AATGGAGGGAATGTCTAG 252
RESULT 64
ABS28814/c
ID ABS28814 standard; DNA; 402 BP.
XX AC ABS28814;
XX DT 25-FEB-2003 (first entry)
XX DE Human liver single exon probe, SEQ ID No 3804.
XX KW Human; single exon nucleic acid probe; liver; cirrhosis;
XX KW hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX KW coronary heart disease; ss.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00609408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-488898/53.
XX DR
XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human adult liver.
XX PS Claim 1; SEQ ID NO 3804; 658pp; English.
XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
XX CC measuring human gene expression in a sample derived from human adult
XX CC liver, comprising one of 13109 defined nucleotide sequences given in the
XX CC specification (or complements/ fragments). The probe hybridises at high
XX CC stringency to a nucleic acid molecule expressed in the human adult liver.
XX CC (1) may be used for predicting, measuring and displaying gene expression
XX CC in samples derived from human adult liver. The genes identified may be
```

```
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABS25011-ABS51005 represent human
CC liver single exon nucleic acid probes of the invention. Note: The
CC sequence information for this patent does not appear in the printed
CC specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 4,58e-31 Length: 402
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.22% Indels: 0
DB: 4 Gaps: 0
US-09-787-097-12 (1-1429) x ABS28814 (1-402)
QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
DB 389 GCTTGCCAATGCAACGGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGAG 330
QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrlGlyAspProThr 1101
DB 329 AACCTGACCACAGGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCACC 270
QY 1102 AsnGlyGlyLysCysGln 1107
DB 269 AATGGAGGGAATGTCTAG 252
RESULT 65
AAI03665/c
ID AAI03665 standard; DNA; 402 BP.
XX AC AAI03665;
XX DT 09-OCT-2001 (first entry)
XX DE Probe #3656 used to measure gene expression in human breast sample.
XX KW Probe; human; breast disease; breast cancer; development disorder; ss;
XX KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.
XX OS Homo sapiens.
XX PN WO200157270-A2.
XX PD 09-AUG-2001.
XX PF 29-JAN-2001; 2001WO-US000661.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00609408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA Penn SG, Hanzel DK, Chen W, Rank DR;
XX PI WPI; 2001-476286/51.
XX DR
XX PT Novel single exon nucleic acid probe used to measuring gene expression in
XX PT a human breast.
XX PS Claim 25; SEQ ID NO 3656; 322pp; English.
XX CC The present invention relates to novel single exon nucleic acid probes.
```

CC The present sequence is one such probe. The probes are useful for  
 CC measuring human gene expression in a human breast sample, where the probe  
 CC hybridises at high stringency to a nucleic acid expressed in the human  
 CC breast. The probes are useful for predicting, diagnosing, grading,  
 CC staging, monitoring and prognosing diseases of the human breast,  
 CC particularly those diseases with polygenic aetiology. The diseases  
 CC include: breast cancer, disorders of development, inflammatory diseases  
 CC of the breast, fibrocystic changes, proliferative breast disease and non-  
 CC carcinoma tumours. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences  
 XX  
 SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4,58e-31 Length: 402  
 Score: 46.00 Matches: 46  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.22% Indels: 0  
 DB: 5 Gaps: 0

US-09-787-097-12 (1-1429) x AA103665 (1-402)

QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081  
 DB 389 GCTTGCATGCAACGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAGTGTGAG 330  
 QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrglyAspProThr 1101  
 DB 329 AACTGACCAACAGCAAGCACTGCGAGACTGTGATATCTGGCTTCTACGGTGTATCCACC 270

QY 1102 AsnGlyGlyLysCysGln 1107  
 DB 269 AATGGAGGGAATGTCTAG 252

RESULT 66

ABS03749/C  
 ID ABS03749 standard; DNA; 402 BP.

AC ABS03749;

XX 19-AUG-2002 (first entry)

DE Human genome-derived single exon probe from lung SEQ ID No 3740.

XX Human; ds; single exon probe; asthma; lung cancer; COPD; ILD;  
 KW chronic obstructive pulmonary disease; interstitial lung disease;  
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;  
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;  
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;  
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;  
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;  
 KW primary ciliary dyskinesia; pulmonary hypertension;  
 KW hyaline membrane disease.

OS Homo sapiens.

XX WQ200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;  
 PI WPI; 2002-114183/15.  
 XX  
 DR Spatially-addressable set of single exon nucleic acid probes, used to  
 PT measure gene expression in human lung samples.  
 PT  
 PS Claim 1; SEQ ID NO 3740; 634bp; English.

XX The invention relates to a spatially-addressable set of single exon  
 CC nucleic acid probes for measuring gene expression in a sample derived  
 CC from human lung comprising single exon nucleic acid probes having one of  
 CC 12614 nucleic acid sequences mentioned in the specification, or their  
 CC complements or the 12387 open reading frames derived from the 12614  
 CC probes. Also included are a microarray comprising the novel set of probes  
 CC; the novel set of probes which hybridise at high stringency to a nucleic  
 CC acid expressed in the human lung; measuring gene expression in a sample  
 CC derived from human lung, comprising (a) contacting the array with a  
 CC collection of detectably labeled nucleic acids derived from human lung  
 CC mRNA; and (b) measuring the label detectably bound to each probe of the  
 CC array; identifying exons in a eukaryotic genome, comprising (a)  
 CC algorithmically predicting at least one exon from genomic sequences of  
 CC the eukaryote; and (b) detecting specific hybridisation of detectably  
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,  
 CC having a fragment identical to the predicted exon, the probe is included  
 CC in the above mentioned microarray; assigning exons to a single gene,  
 CC comprising (a) identifying the expression of each of the exons in several  
 CC tissues and/or cell types using hybridisation to a single exon  
 CC microarrays having a probe with the exon, where a common pattern of  
 CC expression of the exons in the tissues and/or cell types indicates that  
 CC the exons should be assigned to a single gene; a peptide comprising one  
 CC of 12011 sequences, mentioned in the specification, or encoded by the  
 CC probes/open reading frames (ORF). The probes are used for gene expression  
 CC analysis, and for identifying exons in a gene, particularly using human  
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung  
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung  
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,  
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-  
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary  
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,  
 CC Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary  
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The  
 CC present sequence is a single exon probe of the invention. Note: The  
 CC sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 402 BP; 101 A; 94 C; 84 G; 123 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 4,58e-31 Length: 402  
 Score: 46.00 Matches: 46  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 3.22% Indels: 0  
 DB: 6 Gaps: 0

US-09-787-097-12 (1-1429) x ABS03749 (1-402)

QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081

DB 389 GCTTGCATGCAACGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAGTGTGAG 330

QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrglyAspProThr 1101

DB 329 AACTGACCAACAGCAAGCACTGCGAGACTGTGATATCTGGCTTCTACGGTGTATCCACC 270

QY 1102 AsnGlyGlyLysCysGln 1107

DB 269 AATGGAGGGAATGTCTAG 252



```
RESULT 67
AAI23009/c
XX AAI23009 standard; DNA; 137 BP.
XX AC
XX AAI23009;
XX DT
XX 12-OCT-2001 (first entry)
XX DE
XX Probe #12942 for gene expression analysis in human cervical cell sample.
XX DE
XX Probe; human; microarray; gene expression; cervical epithelial cell;
XX KW
XX cervical cancer; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200157278-A2.
XX PD
XX 09-AUG-2001.
XX PF
XX 30-JAN-2001; 2001WO-US000670.
XX PR
XX 04-FEB-2000; 2000US-0180312P.
XX PR
XX 26-MAY-2000; 2000US-0207456P.
XX PR
XX 30-JUN-2000; 2000US-00608408.
XX PR
XX 03-AUG-2000; 2000US-00632366.
XX PR
XX 21-SEP-2000; 2000US-0234587P.
XX PR
XX 27-SEP-2000; 2000US-0236359P.
XX PR
XX 04-OCT-2000; 2000GB-00024263.
XX PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR
XX WPI; 2001-488901/53.
XX PT
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human cervical epithelial cells.
XX PS
XX Claim 25; SEQ ID NO 12942; 487pp; English.
XX CC
XX The present invention relates to human single exon nucleic acid probes
XX (SENPs). The present sequence is one such probe. The SENPs are derived
XX from human HeLa cells. The SENPs can be used to produce a single exon
XX microarray, which can be used for measuring human gene expression in a
XX sample derived from human cervical epithelial cells. By measuring gene
XX expression the probes are therefore useful in grading and/or staging of
XX diseases of the cervix, notably cervical cancer. Note: The sequence data
XX for this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.29e-30 Length: 137
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.15% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAI23009 (1-137)
QY 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
Db 135 TGCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76
QY 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrcGlyAspProThrAsn 1102
Db 75 CTGACCACAGCGCAAGCACTGCGAGACCTGCATATCTGCTTCTACGGTGTATCCCAACAT 16
QY 1103 GlyGlyLysCysGln 1107
Db 15 GGAGGGAATGTCAG 1

US-09-787-097-12 (1-1429) x AAI23009 (1-137)
QY 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
Db 135 TGCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76
QY 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrcGlyAspProThrAsn 1102
Db 75 CTGACCACAGCGCAAGCACTGCGAGACCTGCATATCTGCTTCTACGGTGTATCCCAACAT 16
QY 1103 GlyGlyLysCysGln 1107
Db 15 GGAGGGAATGTCAG 1

RESULT 68
ABA68099/c
XX ABA68099 standard; DNA; 137 BP.
XX AC
XX ABA68099;
XX DT
XX 01-FEB-2002 (first entry)
XX DE
XX Human foetal liver single exon nucleic acid probe #16404.
XX KW
XX Human; foetal liver; gene expression; single exon nucleic acid probe; ss.
XX OS
XX Homo sapiens.
XX PN
XX WO200157277-A2.
XX PD
XX 09-AUG-2001.
XX PF
XX 30-JAN-2001; 2001WO-US000669.
XX PR
XX 04-FEB-2000; 2000US-0180312P.
XX PR
XX 26-MAY-2000; 2000US-0207456P.
XX PR
XX 30-JUN-2000; 2000US-00608408.
XX PR
XX 03-AUG-2000; 2000US-00632366.
XX PR
XX 21-SEP-2000; 2000US-0234587P.
XX PR
XX 27-SEP-2000; 2000US-0236359P.
XX PR
XX 04-OCT-2000; 2000GB-00024263.
XX PA
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PI
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX DR
XX WPI; 2001-483447/52.
XX PT
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human fetal liver.
XX PS
XX Claim 4; SEQ ID NO 16404; 639pp + Sequence Listing; English.
XX CC
XX The invention relates to a single exon nucleic acid probe for measuring
XX human gene expression in a sample derived from human foetal liver. The
XX single exon nucleic acid probes may be used for predicting, measuring and
XX displaying gene expression in samples derived from human fetal liver. The
XX present sequence is a single exon nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX SQ
XX Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.29e-30 Length: 137
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.15% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABA68099 (1-137)
QY 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082
Db 135 TGCCAATGCAACGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76
QY 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrcGlyAspProThrAsn 1102
Db 75 CTGACCACAGCGCAAGCACTGCGAGACCTGCATATCTGCTTCTACGGTGTATCCCAACAT 16
QY 1103 GlyGlyLysCysGln 1107
Db 15 GGAGGGAATGTCAG 1
```

## RESULT 69

AAI48309/C

ID AAI48309 standard; DNA; 137 BP.

XX AC

XX AAI48309;

XX AC

XX 17-OCT-2001 (first entry)

XX DT

XX DE Probe #16995 used to measure gene expression in human placenta sample.

XX XX

XX KW Probe; microarray; human; placenta; antenatal diagnosis;

XX KW genetic disorder; ss.

XX XX

XX OS Homo sapiens.

XX XX

XX PN WO200157272-A2.

XX XX

XX PD 09-AUG-2001.

XX XX

XX PF 30-JAN-2001; 2001WO-US000663.

XX XX

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR

XX PR 30-JUN-2000; 2000US-00608408.

XX PR

XX PR 03-AUG-2000; 2000US-00632366.

XX PR

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR

XX PR 04-OCT-2000; 2000GB-00024263.

XX XX

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX XX

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX

XX WPI; 2001-488897/53.

XX XX

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing

XX PT gene expression in human placenta.

XX XX

XX PS Claim 25; SEQ ID NO 16995; 654pp; English.

XX XX

XX CC The present invention relates to single exon nucleic acid probes (SENP).

XX CC

XX CC The present sequence is one such probe. The probes are useful for

XX CC

XX CC producing a microarray for predicting, measuring and displaying gene

XX CC

XX CC expression in samples derived from human placenta. The probes are useful

XX CC

XX CC for antenatal diagnosis of human genetic disorders

XX XX

XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

XX XX

XX SQ Alignment Scores:

XX XX

XX XX Pred. No.: 1.29e-30 Length: 137

XX XX Score: 45.00 Matches: 45

XX XX Percent Similarity: 100.00% Conservative: 0

XX XX Best Local Similarity: 100.00% Mismatches: 0

XX XX Query Match: 3.15% Indels: 0

XX XX DB: 4 Gaps: 0

XX XX

XX US-09-787-097-12 (1-1429) x AAI48309 (1-137)

XX XX

XX QY 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082

XX Db

XX Db 135 TGCCAATGCAACGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGGAAC 76

XX QY

XX QY 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102

XX Db

XX Db 75 CTGACCAACAGCAAGCACTGCGAGCACTGATATCTGCTTCTACGGTATCCCAACAT 16

XX QY

XX QY 1103 GlyGlyLysCysGln 1107

XX Db

XX Db 15 GGAGGGAATGTCTAG 1

## RESULT 70

ABA50165/c

ABA50165 standard; DNA; 137 BP.

XX AC

XX ABA50165;

XX XX

XX DT 01-FEB-2002 (first entry)

XX XX

XX DE Human breast cell single exon nucleic acid probe #8860.

XX KW

XX KW Human; microarray; single exon probe; gene expression; breast; disease;

XX KW cancer; ss.

XX XX

XX OS Homo sapiens.

XX XX

XX XX WO200157271-A2.

XX XX

XX PD 09-AUG-2001.

XX XX

XX PF 30-JAN-2001; 2001WO-US000662.

XX XX

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR

XX PR 30-JUN-2000; 2000US-00608408.

XX PR

XX PR 03-AUG-2000; 2000US-00632366.

XX PR

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR

XX PR 04-OCT-2000; 2000GB-00024263.

XX XX

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX XX

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX XX

XX WPI; 2001-496933/54.

XX XX

XX PT New spatially-addressable set of single exon nucleic acid probes, useful

XX PT for measuring gene expression in sample derived from human breast,

XX PT comprises number of single exon nucleic acid probes.

XX XX

XX PS Claim 4; SEQ ID NO 8860; 327pp + Sequence Listing; English.

XX XX

The invention relates to a spatially-addressable set of single exon nucleic acid probes for measuring gene expression in a sample derived from human breast and Bt 474 cells. The method involves contacting the probes with a collection of detectably labelled nucleic acids derived from mRNA of human breast, and then measuring the label bound to each probe of the microarray. The probes are useful for verifying the expression of regions of genomic DNA predicted to encode proteins. They are useful for gene discovery, and for determining predisposition and/or prognosing breast disease. Gene expression analysis is useful for assessing the toxicity of chemical agents on cells. The microarray of this invention presents a far greater diversity of probes for measuring gene expression, with far less bias than expressed sequence tag microarrays. The method is suitable for rapid production of functional information from genomic sequence. The present sequence is a single exon nucleic acid probe of the invention. Note: the sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

XX XX

XX SQ Alignment Scores:

XX XX

XX XX Pred. No.: 1.29e-30 Length: 137

XX XX Score: 45.00 Matches: 45

XX XX Percent Similarity: 100.00% Conservative: 0

XX XX Best Local Similarity: 100.00% Mismatches: 0

XX XX Query Match: 3.15% Indels: 0

XX XX DB: 4 Gaps: 0

XX XX

XX US-09-787-097-12 (1-1429) x ABA50165 (1-137)

XX QY

XX QY 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082

XX Db

XX Db 135 TGCCAATGCAACGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAAGTGTGGAAC 76

QY 1083 LeuThrThrClyLysHisCysGluThrCysIleSerClyPheTyrClyAspProThrAsn 1102  
 |||||  
 Db 75 CTGACACAGCAGCACTCGAGACTGCATATCTGGCTTCTACGGTGATCCCAAT 16  
 |||||

QY 1103 GlyClyLysCysGln 1107  
 |||||  
 Db 15 GGAGGGAATGTCAG 1  
 |||||

RESULT 71  
 ABA35129/c  
 ID ABA35129 standard; DNA; 137 BP.  
 XX

AC ABA35129;

XX 23-JAN-2002 (first entry)

XX Probe #13595 for gene expression analysis in human heart cell sample.

XX Human; gene expression; heart; microarray; vascular system; probe;  
 KW cardiovascular disease; hypertension; cardiac arrhythmia;  
 KW congenital heart disease; ss.

XX Homo sapiens.

XX WO200157274-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000666.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488999/53.

XX Single exon nucleic acid probes for analyzing gene expression in human

XX hearts.

XX Claim 4; SEQ ID NO 13595; 530pp; English.

XX The present invention relates to single exon nucleic acid probes for

XX measuring human gene expression in a sample derived from human heart. The

XX present sequence is one such probe. The probes may be used for

XX predicting, measuring and displaying gene expression in samples derived

XX from the human heart via microarrays. By measuring gene expression, the

XX probes are useful for predicting, diagnosing, grading, staging,

XX monitoring and prognosing diseases of the human heart and vascular system

XX e.g. cardiovascular disease, hypertension, cardiac arrhythmias and

XX congenital heart disease. Note: The sequence data for this patent did not

XX form part of the printed specification, but was obtained in electronic

XX format directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

QY 1063 CysGlnCysAsnGlyHisSerIysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082  
 |||||  
 Db 135 TGCCCAATGCACGCCGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAACTGTGAGAAC 76  
 |||||

QY 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102  
 |||||  
 Db 75 CTGACACAGCAGCACTCGAGACTGCATATCTGGCTTCTACGGTGATCCCAAT 16  
 |||||

QY 1103 GlyClyLysCysGln 1107  
 |||||  
 Db 15 GGAGGGAATGTCAG 1  
 |||||

RESULT 72

AAK42239/c

ID AAK42239 standard; DNA; 137 BP.

XX

AC AAK42239;

XX 06-NOV-2001 (first entry)

XX Human bone marrow expressed single exon probe SEQ ID NO: 16796.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KW microarray; cancer; leukaemia; lymphoma; myeloma; ss.

XX Homo sapiens.

XX WO200157276-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US000668.

XX 04-FEB-2000; 2000US-0180312P.

XX 26-MAY-2000; 2000US-0207456P.

XX 30-JUN-2000; 2000US-00608408.

XX 03-AUG-2000; 2000US-00632366.

XX 21-SEP-2000; 2000US-0234687P.

XX 27-SEP-2000; 2000US-0236359P.

XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488999/53.

XX Human genome-derived single exon nucleic acid probes useful for analyzing

XX gene expression in human bone marrow.

XX Example 4; SEQ ID NO 16796; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

XX bone marrow. They can be used to measure gene expression in bone marrow

XX samples, which may enable the improved diagnosis and treatment of cancers

XX such as lymphoma, leukaemia and myeloma. The present sequence is one of

XX the probes of the invention

XX SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.29e-30 Length: 137

Score: 45.00 Matches: 45

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 3.15% Indels: 0

DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAK42239 (1-137)

QY 1063 CysGlnCysAsnGlyHisSerIysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082

```

Db      135 TGCCAATGCACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76
Qy      1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyGlyAspProThrAsn 1102
Db      75 CTGACCACAGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAACAT 16
Qy      1103 GlyGlyLysCysGln 1107
Db      15 GGAGGGAATGTCTAG 1
RESULT 73
AAK16486/c
ID AAK16486 standard; DNA; 137 BP.
XX
AC AAK16486;
XX
XX 05-NOV-2001 (first entry)
XX Human brain expressed single exon probe SEQ ID NO: 16477.
XX Human; brain expressed exon; gene expression analysis; probe; microarray;
XX Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer;
XX ss.
XX Homo sapiens.
XX WO200157275-A2.
XX PN
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
XX brains.
XX Example 4; SEQ ID NO 16477; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
XX probes which are derived from genomic sequences expressed in the human
XX brain. They can be used to measure gene expression in brain cell samples,
XX which may enable the diagnosis and improved treatment of nervous system
XX diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
XX epilepsy and cancers. The present sequence is one of the probes of the
XX invention
XX
SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.29e-30 Length: 137
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.15% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x AAK16486 (1-137)
Qy 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082

```

```

Db      135 TGCCAATGCACGGCCACAGTAATGATCAATCAGAGCATCTGTGAGAAGTGTGAGAAC 76
Qy      1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyGlyAspProThrAsn 1102
Db      75 CTGACCACAGCAAGCACTCGAGACCTGCATATCTGGCTTCTACGGTGATCCCAACAT 16
Qy      1103 GlyGlyLysCysGln 1107
Db      15 GGAGGGAATGTCTAG 1
RESULT 74
ABS41845/c
ID ABS41845 standard; DNA; 137 BP.
XX
XX ABS41845;
XX
XX 25-FEB-2003 (first entry)
XX Human liver single exon probe, SEQ ID NO 16835.
XX Human; single exon nucleic acid probe; liver; cirrhosis;
XX hyperlipoproteinaemia; hyperlipidaemia; hypercholesterolaemia;
XX coronary heart disease; ss.
XX Homo sapiens.
XX OS
XX WO200157273-A2.
XX PN
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000664.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX PA
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488898/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
XX gene expression in human adult liver.
XX Claim 4; SEQ ID NO 16835; 658pp; English.
XX
XX The invention relates to a single exon nucleic acid probe (SEN) (I) for
XX measuring human gene expression in a sample derived from human adult
XX liver, comprising one of 13109 defined nucleotide sequences given in the
XX specification (or complements/ fragments). The probe hybridises at high
XX stringency to a nucleic acid molecule expressed in the human adult liver.
XX (I) may be used for predicting, measuring and displaying gene expression
XX in samples derived from human adult liver. The genes identified may be
XX involved in genetic liver diseases such as cirrhosis,
XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
XX associated with coronary heart disease. ABS25011-ABS1005 represent human
XX liver single exon nucleic acid probes of the invention. Note: The
XX sequence information for this patent does not appear in the printed
XX specification but was obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pat_sequences
XX
SQ Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1.29e-30 Length: 137
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0

```

Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.15% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x ABS41845 (1-137)

QY 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082  
Db 135 TGCCAATGCAAGCCACAGTAATGCAATCATCAGAGCATCTGTGAGAGTGTGAGAAC 76

QY 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTy-GlyAspProThrAsn 1102  
Db 75 CTGACCACAGGAAGCACTGCCAGACCTGCATATCTGGCTTCTACGGTGATCCCAACAT 16

QY 1103 GlyGlyLysCysGln 1107  
Db 15 GGAGGGAATGTCAG 1

RESULT 75  
AAI08671/c  
ID AAI08671 standard; DNA; 137 BP.  
XX  
AC AAI08671;  
XX  
DT 09-OCT-2001 (first entry)  
XX  
DE DE Probe #8662 used to measure gene expression in human breast sample.  
XX  
KW Probe; human; breast disease; breast cancer; development disorder; ss;  
KW inflammatory disease; proliferative breast disease; non-carcinoma tumour.  
XX  
OS Homo sapiens.  
XX  
PN WO200157270-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 29-JAN-2001; 2001WO-US000661.  
XX  
PR 04-FEB-2000; 2000US-0180312P.  
PR 26-MAY-2000; 2000US-0207456P.  
PR 30-JUN-2000; 2000US-00608408.  
PR 03-AUG-2000; 2000US-00632366.  
PR 21-SEP-2000; 2000US-0234587P.  
PR 27-SEP-2000; 2000US-0236359P.  
PR 04-OCT-2000; 2000GE-00024263.  
XX  
PA (MOLE-) MOLECULAR DYNAMICS INC.  
XX  
PI Penn SG, Hanzel DK, Chen W, Rank DR;  
XX  
DR WPI; 2001-476286/51.  
XX  
PT Novel single exon nucleic acid probe used to measuring gene expression in  
PT a human breast.  
XX  
PS Claim 25; SEQ ID NO 8662; 322pp; English.  
XX

CC The present invention relates to novel single exon nucleic acid probes.  
CC The present sequence is one such probe. The probes are useful for  
CC measuring human gene expression in a human breast sample, where the probe  
CC hybridises at high stringency to a nucleic acid expressed in the human  
CC breast. The probes are useful for predicting, diagnosing, grading,  
CC staging, monitoring and prognosing diseases of the human breast,  
CC particularly those diseases with polygenic aetiology. The diseases  
CC include: breast cancer, disorders of development, inflammatory diseases  
CC of the breast, fibrocystic changes, proliferative breast disease and non-  
CC carcinoma tumours. Note: The sequence data for this patent did not form  
CC part of the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences

XX Sequence 137 BP; 26 A; 35 C; 35 G; 41 T; 0 U; 0 Other;  
SQ

Alignment Scores:  
Pred. No.: 1-29e-30 Length: 137  
Score: 45.00 Matches: 45  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.15% Indels: 0  
DB: 5 Gaps: 0

US-09-787-097-12 (1-1429) x AAI08671 (1-137)

QY 1063 CysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsn 1082  
Db 135 TGCCAATGCAAGCCACAGTAATGCAATCATCAGAGCATCTGTGAGAGTGTGAGAAC 76

QY 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTy-GlyAspProThrAsn 1102  
Db 75 CTGACCACAGGAAGCACTGCCAGACCTGCATATCTGGCTTCTACGGTGATCCCAACAT 16

QY 1103 GlyGlyLysCysGln 1107  
Db 15 GGAGGGAATGTCAG 1

Search completed: March 1, 2004, 07:55:33  
Job time : 1707 secs

THIS PAGE BLANK (USPIO)

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 1, 2004, 07:35:35 ; Search time 1026 Seconds  
(without alignments)

5024.435 Million cell updates/sec

Title: US-09-787-097-12

Perfect score: 1429

Sequence: 1 MVAAAAAATEARLRRTAATA.....KSGAVRNKQPPAPQPTCI 1429

Scoring table:

Xgapop 60.0	Xgapext 60.0
Ygapop 60.0	Ygapext 60.0
Fgapop 6.0	Fgapext 7.0
Delop 6.0	Delext 7.0

Searched: 2353733 seqs, 1803733377 residues

Word size: 1

Total number of hits satisfying chosen parameters: 4700994

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame\_p2n.model -DEV=xlp  
-DB=/cn2\_1/USPTO\_pool\_p/US09787097/runat\_26022004\_063017\_23865/app\_query.fasta\_1.1607  
-DB=Published Applications NA -GPMT=fastap -SUFFIX=oligo.rnpb -MINMATCH=0.1  
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=oligo  
-TRANS=human40.cdi -LIST=100 -DOCLIGN=200 -THR SCORE=quality -THR MIN=1  
-ALIGN=75 -MODE=LOCAL -OUTFMT=ptco -NORM=ext -HEARSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US09787097 -CGN 1 1 136 -runat 26022004\_063017\_23865  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=60 -XGAPEXT=60 -DELOP=6 -DELEXT=7

Database :

Published Applications NA:

1:	/cn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2:	/cn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
3:	/cn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
4:	/cn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
5:	/cn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
6:	/cn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
7:	/cn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8:	/cn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9:	/cn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10:	/cn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11:	/cn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12:	/cn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13:	/cn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
14:	/cn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
15:	/cn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
16:	/cn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
17:	/cn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
18:	/cn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1141	79.8	8589	9	US-09-893-238-14	Sequence 14, Appl
2	979	68.5	4072	9	US-09-893-238-16	Sequence 16, Appl
3	503	35.2	6370	9	US-09-893-238-12	Sequence 12, Appl
4	496	34.7	2625	9	US-09-893-238-18	Sequence 18, Appl
5	206	14.4	8827	9	US-09-893-238-1	Sequence 1, Appl
6	94	6.6	2419	9	US-09-893-238-8	Sequence 8, Appl
c	7	6.4	625	14	US-10-198-846-5869	Sequence 5869, Ap
c	8	5.7	246	9	US-09-736-457-966	Sequence 966, App
c	9	5.7	246	9	US-09-902-941-966	Sequence 966, App
c	10	5.7	246	9	US-09-849-626-966	Sequence 966, App
c	11	82	5.7	246	US-10-017-754-966	Sequence 966, App
c	12	82	5.7	246	US-10-113-872-966	Sequence 5, Appl
c	13	77	5.4	207433	US-10-277-216-5	Sequence 5, Appl
c	14	77	5.4	207433	US-10-126-022-5	Sequence 10, Appl
c	15	76	5.3	1051	US-09-893-238-10	Sequence 5, Appl
c	16	73	5.1	90050	US-09-893-238-5	Sequence 5, Appl
c	17	69	4.8	531	US-09-864-761-15900	Sequence 15900, A
c	18	64	4.5	193	US-09-864-761-32404	Sequence 32404, A
c	19	60	4.2	182	US-09-864-761-21460	Sequence 21460, A
c	20	57	4.0	375	US-09-864-761-4720	Sequence 4720, Ap
c	21	51	3.6	226	US-09-893-238-6	Sequence 6, Appl
c	22	46	3.2	212	US-09-796-692-7610	Sequence 7610, Ap
c	23	46	3.2	212	US-10-040-862-7610	Sequence 7610, Ap
c	24	46	3.2	212	US-10-057-475B-7610	Sequence 7610, Ap
c	25	46	3.2	212	US-10-154-884B-7610	Sequence 7610, Ap
c	26	45	3.2	402	US-09-864-761-3581	Sequence 3681, Ap
c	27	45	3.1	137	US-09-864-761-20449	Sequence 20449, A
c	28	45	3.1	221	US-09-893-238-7	Sequence 7, Appl
c	29	33	2.3	5973	US-09-893-238-4	Sequence 4, Appl
c	30	30	2.1	17056	US-09-893-238-3	Sequence 3, Appl
c	31	29	2.0	6733	US-10-197-824-1	Sequence 1, Appl
c	32	28	2.0	594	US-10-027-632-264639	Sequence 264639, A
c	33	28	2.0	594	US-10-027-632-264640	Sequence 264640, A
c	34	26	1.8	629	US-10-027-632-270240	Sequence 270240, A
c	35	22	1.5	415	US-09-893-381-1946	Sequence 1916, Ap
c	36	22	1.5	5312	US-10-311-453-170	Sequence 170, App
c	37	20	1.4	60	US-09-908-975-15421	Sequence 15421, A
c	38	18	1.3	65	US-09-908-975-29990	Sequence 29990, A
c	39	16	1.1	233	US-09-728-446-450	Sequence 450, App
c	40	14	1.0	2572	US-09-925-300-486	Sequence 486, App
c	41	14	1.0	3897	US-10-198-846-13651	Sequence 13651, A
c	42	14	1.0	3901	US-10-240-423-1425	Sequence 1425, Ap
c	43	14	1.0	3901	US-10-177-293-440	Sequence 440, App
c	44	14	1.0	3901	US-10-465-572-17	Sequence 17, Appl
c	45	14	1.0	3901	US-10-159-563-292	Sequence 292, App
c	46	14	1.0	3901	US-10-372-683-40	Sequence 40, Appl
c	47	14	1.0	225883	US-10-175-523-57	Sequence 57, Appl
c	48	13	0.9	105	US-10-029-386-14368	Sequence 14368, A
c	49	13	0.9	520	US-10-027-632-181373	Sequence 181373, A
c	50	13	0.9	575	US-10-029-386-664	Sequence 664, App
c	51	13	0.9	935	US-10-027-632-155757	Sequence 155757, A
c	52	13	0.9	2257	US-10-108-260A-876	Sequence 876, App
c	53	13	0.9	2656	US-09-965-529-39	Sequence 39, Appl
c	54	13	0.9	2656	US-09-969-680A-39	Sequence 39, Appl
c	55	13	0.9	6691	US-10-301-822-3	Sequence 3, Appl
c	56	13	0.9	8772	US-09-788-711A-3	Sequence 3, Appl
c	57	13	0.9	8871	US-09-788-711A-1	Sequence 21, Appl
c	58	13	0.9	9321	US-10-311-623-21	Sequence 523, App
c	59	13	0.9	10531	US-10-225-567A-523	Sequence 931, App
c	60	13	0.9	24370	US-10-292-798-931	Sequence 31, Appl
c	61	13	0.9	68233	US-10-034-650-31	Sequence 31, Appl
c	62	12	0.8	287	US-09-864-761-24815	Sequence 24815, A
c	63	12	0.8	361	US-10-027-632-28416	Sequence 28416, A
c	64	12	0.8	394	US-10-198-846-13018	Sequence 13018, A
c	65	12	0.8	411	US-10-040-739-599	Sequence 599, App
c	66	12	0.8	436	US-09-864-761-2885	Sequence 2885, App
c	67	12	0.8	455	US-10-198-846-5901	Sequence 5901, Ap
c	68	12	0.8	465	US-10-09-330-213-151	Sequence 151, App
c	69	12	0.8	468	US-10-027-632-108988	Sequence 108988, A
c	70	12	0.8	526	US-10-029-386-11510	Sequence 11510, A
c	71	12	0.8	575	US-09-864-761-8070	Sequence 8070, App
c	72	12	0.8	641	US-10-101-464A-230	Sequence 230, App
c	73	12	0.8	659	US-10-029-386-20268	Sequence 20268, A





QY 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSerServ 506  
Db 1454 CATGGAGTATATTACACACCCAGGGTCCCTTTGTGCAAGGGGGTTACGGCCATACAGTG 1513  
QY 506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrIlysalapheserAlaA 526  
Db 1514 TTACACCATAGGACCGGGCCCTATACGTTTCATGGTGGCTACAGGCTTTTCAGTGCCA 1573  
QY 526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTyrThrI 546  
Db 1574 ATAAGTACCGGGTTGCGATGATCTCTACCGATATGATGGATACCCAGATGTGGACCA 1633  
QY 546 leLeuLysAspSerArgPheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566  
Db 1634 TTCTTAAGACACCGCCGATTTTCCGTTACTTCACACAGCTGTGATAGTGAAGCCA 1693  
QY 566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586  
Db 1694 TCGTGGTGTGGGGGAAACACACACATGACATCTATGAGCCATGGCGCAATGCT 1753  
QY 586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606  
Db 1754 TCTCTTCAGATTTCATGGCCATGACATTCGCTGTGACCGCTGGTCAGTGTCTTCCAGAC 1813  
QY 606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValIleHisAsnSerThrM 626  
Db 1814 CTGATCTCCACCATGATGTCACAGATTGGCCATTCACAGCTTTTACACACAGACCA 1873  
QY 626 etTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerG 646  
Db 1874 TGTATGTGTGGTGGTTTCAATAGTCTCTCTCCAGCAGATCCTGGTATTTCACCTGG 1933  
QY 646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
Db 1934 AACAGTGTGATCGCATCGAGTGAAGCCGCTGTTTAGCAGCAGGACCTGTGTATTCGGT 1993  
QY 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTyrAlaLeuAlaThrAspGluGlnG 686  
Db 1994 GTGTGTGGAAACACAGAGGTGCTCTCAGTGTATCTCGTGGCGGTGGCACTGTATGACACAG 2053  
QY 686 luGluLysLeuLysSerGluCysPheSerIlyArgThrLeuAspHisAspArgCysAspG 706  
Db 2054 AAGAAAAGTTAAATCAGAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113  
QY 706 leHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTyrCysAsnAspH 726  
Db 2114 AGCACACAGATTGTTACAGCTGCACAGCCAAACCAATGACTGCCACTGGTGCATGACC 2173  
QY 726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746  
Db 2174 ATTGTGTCCCGAAGAACCCAGCTGCTCAGAGGCCAGATCTCCATTTTAGGTATGAGA 2233  
QY 746 snCysProLysAspAsnProMetTyrTyrCysAsnLysIlyThrSerCysArgSerCysA 766  
Db 2234 ATTGCCCCAGGATAACCCATATGCTACTGTAAACAGAGACCCAGCTGCAGAGCTGTG 2293  
QY 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786  
Db 2294 CCCTGACCCAGAACTGCCAGTGGAGGCCCGGATCAGAGTGCATTTGCCCTGCCGAAA 2353  
QY 786 snIleCysGlyIleGlyTyrPheHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806  
Db 2354 ATATCTGTGCAATGGCTGGCAATTTGGTTGGAACCTCATGTTTGAATAATTAATGCAATCTC 2413  
QY 806 ysGluAsnTyrAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826  
Db 2414 AGGAGAAATATGACAAATGCTAAATTTCTGTAGGAACACCAATGCCCTTTTGGCTTCTC 2473  
QY 826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846  
Db 2474 TTACACCCAGAGAAGGTAGAAATTTGCTTAAAGCAGCTGGCAATTAATGCAATCTC 2533

QY 846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866  
Db 2534 AGAGCATGTCCAGCTCACCTTAACCCATGGTTCGGCTTCGGAGATCAATGTGTCT 2593  
QY 866 yrTrpCysTyrGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886  
Db 2594 ACTGTCTCTGGAGATATGTCCCATTTACAATAGTTTACAGTGGATGCCGCTG 2653  
QY 886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906  
Db 2654 AGCCCACTGATGCTGGATTTCTGTGAATTTTATCAGAACCCAGTACTCGGGGACCTGAAG 2713  
QY 906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926  
Db 2714 CTGCAACTGTCATCAACCCACTCAATGATGTCTGTGAAGCCCTGCAACACACAGTG 2773  
QY 926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946  
Db 2774 CTAAAGCAGTCCGACACCATGTCCTTTGAGGACAGCATGTGGAGATTGACCAGCGCA 2833  
QY 946 erSerGluCysMetTyrCysSerAsnMetLysGlnCysValAspSerAsnAlaTyrValA 966  
Db 2834 GCTCTGATGTCATGTGTGAGCAACATGAAGCAGTGTGTGGACTCCCAATGCCCTATGTG 2893  
QY 966 laSerPheProPheGlyGlnCysMetGluTyrTrpThrMetSerThrCysProGluA 986  
Db 2894 CCTCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCACCTGCCCTCTGAAA 2953  
QY 986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTyrCyst 1006  
Db 2954 ATTGTTCCAGGCTACTGTACTGTAGTCTGTAGTCTTTGGAGCAACCCAGGCTGTGGCTGTGTA 3013  
QY 1006 hrAspProSerAsnThrGlyLysGlyCysIleGluGlySerTyrLysGlyProValL 1026  
Db 3014 CTGATCCAGCAATCTGGCAAGGAAATGCATAGAGGTTCTCTATAAGGACCATGTA 3073  
QY 1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046  
Db 3074 AGATGCTTCGCAAGCCCTACAGAAATTTCTATCCACAGCCCTGCTCAATTCACGCA 3133  
QY 1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066  
Db 3134 TGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCATTCTCATCTGTCAGTTGCCAATGCA 3193  
QY 1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086  
Db 3194 ACGGCCACAGTAATGCATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCACAG 3253  
QY 1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyLysC 1106  
Db 3254 GCAAGCCTGCGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAATGGAGGGAAT 3313  
QY 1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126  
Db 3314 GTCAGCCATGCAAGTGCATGGGCCCGGCTCTGTGCAACACCAACACGSGCAGTGTCT 3373  
QY 1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146  
Db 3374 TCTGACCACCAAGGGCGTCAAGGGGACGAGTGCAGCTATGTGAGTAGAATAATCGAT 3433  
QY 1146 yrGlnGlyAsnProIleuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166  
Db 3434 ACCAAGGAACCCCTCTCAGAGGAACATGTTATATATCTCTTCTTATGACTATCAGTTCA 3493  
QY 1166 hrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186  
Db 3494 CCTTTAGTCTATCCAGGAAGATGATCGCTATTACACAGCTATCAATTTTGTGGCTACTC 3553  
QY 1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206  
Db 3554 CTGACGAACAAACAGGAGATTGGACATGTTTCATCAATGCCCTCCAGAAATTTCAACCTCA 3613  
QY 1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluMetProValV 1226

Db 3614 ACATCACTGGCTGGCAGTTTCTAGCTGGAACCCAGGCTGGAGAGAGATGCTGTGTTG 3673  
Qy 1226 alSerlystrAnleLysGlyLeuTyrlsAspSerPheSerAsnGluLysPheAspPheA 1246  
Db 3674 TTTCAAAACCAACATTAAGGAGTACAAAGATAGTTCTCTAATGAGAGATTTGATTTTC 3733  
Qy 1246 rGAsnHisProAsnleLysPheValTyrlsSerAsnPheThrTrpProleLysI 1266  
Db 3734 GCAACCAACCAATATACATTTCTTTGTTATCTAGTAAATTTCACTGGCCCAACAA 3793  
Qy 1266 leGlnleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrP 1286  
Db 3794 TTCAGATGGCTTCTCTCAGCAGACAAATTTATGACCTGGTACAGTTCTTCGTGACTT 3853  
Qy 1286 hePheSerCysPheLeuSerLeuLeuValAlaAlaValValTlPlylLysGlns 1306  
Db 3854 TCTTCAGTTGTTCTCTCTTCTCTCTGTTGCTGCTGTTGTTGGAAGATCAACAA 3913  
Qy 1306 erCysTrpAlaSerA:GArgGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerA 1326  
Db 3914 GTTGTGGGCTCCTCAGACGCTAGAGCAACTTCTTCAGAGATGCAACAGATGGCCAGCC 3973  
Qy 1326 rGProPheAlaSerValAsnValAlaLeuGluThrAspGluGluProProAspLeuLeG 1346  
Db 3974 GTCCCTTTGCTCTGTAATGTCGCTTGGAAACAGATGAGGAGCTTCTGATCTTATTG 4033  
Qy 1346 lyGlySerileLysThrValProLysProileAlaLeuGluProCysPheGlyAsnLysA 1366  
Db 4034 GGGGGAGTATAAGACTGTTCCTCAACACCATTCAGCTGGAGCGGTCTTTTGGCAACAA 4093  
Qy 1366 lalaValLeuSerValPheValArgLeuProArgGlyLeuGlylleProProProG 1386  
Db 4094 CCCTGTCTCTCTGTGTGTGTGAGGCTCCTCTGAGGCTGGGTGGCATCCCTCTCTCTG 4153  
Qy 1386 lyGlnSerGlyLeuAlaValAlaSerAlaLeuValAspilleSerGlnGlnMetProileV 1406  
Db 4154 GGAGTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4213  
Qy 1406 alTyrlsGluLysSerGlyAlaValArgAsnArgLysGlnGlnProProAlaGlnProG 1426  
Db 4214 TGTACAGGAGAGAGTCAAGAGCGCTGAGAAACCGAAGCAGCAGGCCCCCTGACAGCCTG 4273  
Qy 1426 lyThrCysIle 1429  
Db 4274 GGACCTGCATC 4284

## RESULT 2

US-09-893-238-16  
; Sequence 16, Application US/09893238  
; Patient No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Moore, K.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893, 238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245, 041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093, 630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104, 978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 16  
; LENGTH: 4072  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-893-238-16

Alignment Scores:  
Pred. No.: 0 Length: 4072  
Score: 979.00 Matches: 1179  
Percent Similarity: 99.66% Conservative: 0  
Best Local Similarity: 99.66% Mismatches: 2  
Query Match: 68.51% Indels: 4  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-893-238-16 (1-4072)

Qy 87 AlaAlaAlaValSerGlySerAlaAlaAlaAlaLysGluCysAspArgProCysVal 106  
Db 256 CGCGCGCGGTGTCGGCTCGCGCGCAGCCGAGGCCAAGGAATGTACCCGCCGTGTGTC 315  
Qy 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTTPValGly 126  
Db 316 AACGGCGGTGCTGCAACCTGGCCACCGCCAGTGGCTGTGCCCGCGCTGGTGGGC 375  
Qy 127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146  
Db 376 GAGCAATGCCAGCACTCGCGGGCGCTTCAGACTAACTGGATCTTCTGGGTGTGTGACA 435  
Qy 147 AspGlyProGlyAsnTyrlsTyrlsTyrlsCysThrTrpLeuLeuGluGlyGlnPro 166  
Db 436 CATGGACTGGAAATTAATAATCAAAACGAAGTGCAGTGGCTCATTTGAAGACAGCCA 495  
Qy 167 AsnArgIleMetArgLeuA:GpHeAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186  
Db 496 AATAAGATAATAGACTTCGTTTCAATCATTTTGTACAGAGTGTAGTTGGGACCAATTA 555  
Qy 187 TyrlsValTyrlsAspGlyAspSerileTyrlsAlaProLeuValAlaAlaPheSerGlyLeu 206  
Db 556 TATGTTTATGATGGGACTCAATTTATGACCGCTAGTGTGCTGATTTAGTGGCTCAT 615  
Qy 207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrlsAla 226  
Db 616 GTTCTCAGAGAGATGCAATAGACTGTCCCTGAGGTGTTGTCACATCAGCTTATGCC 675  
Qy 227 LeuLeuHisPhePheSerAspAlaAlaTyrlsAsnLeuThrGlyPheAsnleThrTyrlsSer 246  
Db 676 TTGCTGATATTTTGTAGTGTGCTTAATAATTGACTGGATTTAATATATTCTTACAGT 735  
Qy 247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266  
Db 736 TTTGATATGTCCTCAATTAATCTCTCAGCGCGAGGAGGTGAAGTCACTAATAGCAGC 795  
Qy 267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286  
Db 796 GA-TACTGTTGAATGTGAATGTTCTGAAAACCTGGAAAGGTGAAGCATGTGCACATTCCTCA 854  
Qy 286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgG 306  
Db 855 CTGTACAGACACTGTGTTTCTCTCATCCAGGCACTGCAATTCAGTCAATGTCAGAGG 914  
Qy 306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnG 326  
Db 915 ATGCTCCTGCTTCTCAGACTGGCAGGCTGCTGGATGTTTCAAGTTCCTGTACAGCTAACCA 974  
Qy 326 nSerPheTrpThrArgGluGluTyrlsSerAsnLeuLysLeuProArgAlaSerHisLysAl 346  
Db 975 GTCAATTTGGACTCGAGAGGAATATTCTAATTAAGTCTCCAGAGCATCTCATTAAGC 1034  
Qy 346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrlsMetPheAsnHisSerAspTy 366  
Db 1035 TGTGCTCAATGGAAACATTTATGTTGGGTGTTGAGGATATATGTTCAACCACTCAGATTA 1094  
Qy 366 rAsnMetValLeuAlaTyrlsAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386  
Db 1095 TAACTGGTCTAGCGTATGACCTTGTCTTAGGAGTGGCTTCCACTTAACCGTCTGT 1154  
Qy 386 lAsnAsnValValValArgTyrlsGlyHisSerLeuAlaLeuTyrlsAspLysIleTyrlsMe 406  
Db 1155 GAACAATGTGTTGTAGATATGTCATCTTTTGGCAATTATACAAGCATAAATTTACAT 1214

Qy 406 tTyrGlyIyIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426  
Db 1215 GTATGGAGGAAAAATTGATTC-AACTGGGAATGTGACCAATGAGTTTGAGATTTTCACA 1273  
Qy 426 leHisAsnGluSerTrpValLeuLeuThrProIyAlaIyGluGlnTrpAlaValValG 446  
Db 1274 TTCAATAATGAGTCATGGGTGTGTGTGACCCCTTAAGGCAAGGAGCAGTATGACAGTGGTTG 1333  
Qy 446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466  
Db 1334 GGCACCTCTCCACACATTTGATACCTGAGAAATGGCGAGTGTCAATGGTGCATCTTTG 1393  
Qy 466 lyHisCysProLeuTrpGlyIyIleSerAsnValGlnGluTrpAspLeuAspLysAsnT 486  
Db 1394 GTCACTGCCCTCTCTATGATATATAAGCAATGTGCAGGAATATGATTTGGATAAGAAC 1453  
Qy 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyIyTrpGlyHisSerSerV 506  
Db 1454 CATGGAGTATATTACACACCCAGGGTGGCCCTTGTGCAAGATGGCGAGTGTACGGCCATAGCAGTG 1513  
Qy 506 aIlyrAspHisArgThrArgAlaLeuTrpValHisGlyGlyTrpLysAlaPheSerAlaA 526  
Db 1514 TTTACGACCATAGGACACGGGCCCTATACGTTTCATGGTGGCTACAAAGGCTTTCAGTGCCA 1573  
Qy 526 snLysTrpArgLeuAlaAspAspLeuTrpArgTrpAspValAspThrGlnMetTrpThrI 546  
Db 1574 ATAAGTACCAGGCTGTCAGATGATCTCTACCGATATGATGTGATACCCAGATGTGGACCA 1633  
Qy 546 leLeuLysAspSerArgPhePheArgTrpLeuHisThrAlaValIleValSerGlyThrM 566  
Db 1634 TTCCTTAAGACACCGCATTTTCCGTTACTTGCACACACAGCTGTGATGTAGTGGAGAACCA 1693  
Qy 566 etLeuValPheGlyGlyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586  
Db 1694 TGTGTGTGTGGGGGAAAAACACACAATGACACATCTATGAGCCATGGCGCAATGCT 1753  
Qy 586 heSerSerAspPheMetAlaTrpAspIleAlaCysAspArgTrpSerValLeuProArgP 606  
Db 1754 TCTCTTCATATTTCATGGGCTATGACANTGCTGTGACCGCTGGTCAGTGTCTCCAGAC 1813  
Qy 606 roAspLeuHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626  
Db 1814 CTGATCTCCACCATGATGTCACAGATTTGGCAATTCACAGCTTTTACACACAGACCA 1873  
Qy 626 etTrpValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerG 646  
Db 1874 TGTATGTGTTCGGTGTGTTTCAATAGTCTCTCTCAGCGACATCCCTGGTATTACCTCGG 1933  
Qy 646 luGlnCysAspAlaHisAspSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
Db 1934 AACAGTGTGATGGCATCGAGTGAAGCCGCTTGTATTAGCAGCAGGACCTGGTATTCCGT 1993  
Qy 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686  
Db 1994 GTGTGTGGAACACAGGGTGTCTCAGTGTATCTCGTGGCGCTGGCAACTGATGAACAG 2053  
Qy 686 luGluLysIleLysSerGluCysPheSerIyAspArgThrLeuAspHisAspArgCysAspG 706  
Db 2054 AAGAAAAGTTAAATCAGAAATGTTTTTCCAAAAAGAACTCTTGACCAATGACAGATGGACC 2113  
Qy 706 lnHisThrAspCysTrpSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726  
Db 2114 AGCACACAGATTGTATACAGCTGCACAGCCCAACCAATGACTGCCACTGGTGCATGACC 2173  
Qy 726 isCysValProArgAsnHisSerCysSerGluGlnIleSerIlePheArgTrpGluA 746  
Db 2174 ATTGTGTCCCCAGAAACACACAGTGTCTCAGAAAGCCAGATCTCTCAATTTTATGATGAGA 2233  
Qy 746 snCysProLysAspAsnProMetTrpTrpCysAsnLysLysThrSerCysArgSerCysA 766  
Db 2234 ATTGCCCAAGGATACCCCTATGTACTGTAAACAGAGACACACAGCTGCAGGAGCTGTG 2293

Qy 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786  
Db 2294 CCCTGGACCAAAATGTCAGTGGAGAGCCCGGAATCAGGAGTGCATTCCTCCCGGAAA 2353  
Qy 786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806  
Db 2354 ATATCTGTGCAATGGCTGGCAATTTGGTGGAAACTCATGTTTGAAAAATTACTACTGCCA 2413  
Qy 806 ysGluAsnTrpAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuLeuAlaSerL 826  
Db 2414 AGGAGAAATTATGACAATGTAAATTTCTGTAGGAACCAACAATGCCCTTTTGGCTTCTC 2473  
Qy 826 eufhrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846  
Db 2474 TTCAACCCAGAGAAAGGTAGAATTTGTCTTAAGCAGCTGCGAATAATCACTCATCTC 2533  
Qy 846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866  
Db 2534 AGAGCATGTCCNAGCTCACCTTAACCCCATGGGTGGGCTTCGGAAGATCAATGTGTCTC 2593  
Qy 866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886  
Db 2594 ACTGTCTGTGGAGATATGTCCCATTTTACAATAGTTTTACTACAGTGGATGCCGCTG 2653  
Qy 886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906  
Db 2654 AGCCAGTGTGTGGATTTCTGTGAATTTTATCAGAACCCAGTACTCGGGACTGAAGG 2713  
Qy 906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926  
Db 2714 CTGCAACTGTCATCAACCCACTCAATGGTAGTGTCTGTGAAGGCTTCGAAACCCACAGTG 2773  
Qy 926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946  
Db 2774 CTAAGCAGTGGCGACACCATGTGCTTGAGGACAGCATGTGGAGATTGCACACAGCGCA 2833  
Qy 946 erSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTrpValA 966  
Db 2834 GCTCTGAGTGCATGTGGTGAGCAACATGAAGCAGTGTGTGGACTCCAAATGCCTATGTGG 2893  
Qy 966 laSerPheProPheGlyGlnCysMetGluTrpTrpThrMetSerThrCysProProGluA 986  
Db 2894 CCTCCTTCCCTTTTGGCCAGTGTATGGAATGGTATACGATGAGCAGCTGCCCCCTGAAA 2953  
Qy 986 snCysSerGlyTrpCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006  
Db 2954 ATTGTTCAGGCTACTGTACCTGTAGTCTATTGCTTGGAGCAACAGGCTGTGGCTGGTGA 3013  
Qy 1006 hrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySerTrpLysGlyProValL 1026  
Db 3014 CTGATCCAGCAATACTGGCAAGGAAATGATAGAGGGTTCTTATAAAGGACACAGTGA 3073  
Qy 1026 ysMetProSerGlnAlaProThrGlyAsnPheTrpProGlnProLeuLeuAsnSerSerM 1046  
Db 3074 AGATGGCTTCCCAAGCCCTTACAGAAATTTCTATCCACAGCCCTGCTCAATTTCCAGCA 3133  
Qy 1046 etCysLeuGluAspSerArgTrpAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066  
Db 3134 TGTGTCTAGAGCAGACAGATACAACTGGTCTTTCACTTCACTGTCCAGCTTGCCAAATGCA 3193  
Qy 1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrG 1086  
Db 3194 ACGGCCACAGTAAATGCAATCAATCAGAGCATCTGTGAGAGTGTGAGAACCTGACCAAG 3253  
Qy 1086 lyLysHisCysGluThrCysIleSerGlyPheTrpGlyAspProThrAsnGlyGlyLysC 1106  
Db 3254 GCAGCACTGGGACACCTGCATATCTGGCTTCTACGGTGATCCCCCAATGAGGGAAAT 3313  
Qy 1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126  
Db 3314 GTACCCATGCAAGTCAATGGGACCGCTCTCTGTGTGCAACACCAACACGCGCAAGTGCT 3373  
Qy 1126 heCysThrThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146



QY 1313 gGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnVa 1333  
Db 1706 AGAGCAACTCTTTCGAGAGATGCAACAGATGGCCAGCCGTCCTTTGGCTCTGTAAATGT 1765  
QY 1333 lAlaLeuGluThrAspGluGluProProAspLeuileGlyGlySerIleLysThrValPr 1353  
Db 1766 CGCTCTGGAAACAGATGAGGAGCTCTGTATCTTATTTGGGGAGATATAAGACTGTCTCC 1825  
QY 1353 oLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheVa 1373  
Db 1826 CAACACCAATTCACCTGGAGCGGTGTTTGGCAACAAAGCCGTCTCTCTGTGTGTGT 1885  
QY 1373 lArgLeuProArgGlyGlyGlyIleProProProGlyGlnSerGlyLeuAlaValAl 1393  
Db 1886 GAGGCTCCCTCGAGGCTGGGTGGCATCCCTCTCTGGGAGTCAAGTCTTGTGTGTGTGT 1945  
QY 1393 aSerAlaLeuValAspIleSerGlnGlnMetProIleValIleValIleValIleVal 1413  
Db 1946 CAGCGCCCTGGTGACATTTCTCAGCAGATGCCGATAGTACAGGAGAGTCAAGGAGC 2005  
QY 1413 aValArgAsnArgLysGlnGlnProProAlaGlnProGlyThrCysIle 1429  
Db 2006 CFTGAGAAACCGGAGCAGCAGCCCTCTGCACAGCCTGGGACCTGTCATC 2054

## RESULT 4

US-09-893-238-18  
; Sequence 18, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-893-238-18

Alignment Scores:  
Pred. No.: 0 Length: 2625  
Score: 496.00 Matches: 696  
Percent Similarity: 99.43% Conservative: 0  
Best Local Similarity: 99.43% Mismatches: 2  
Query Match: 34.71% Indels: 4  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-893-238-18 (1-2625)

QY 87 AlaAlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
Db 256 GCGGCGCGGTGTCGCGCTCGGCGCGAGCCGAGCCAGGAAATGACCGCGCGTGTGTC 315  
QY 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126  
Db 316 AACGGCGGTGCTGCAACCTTGGCACCGGCGAGTGGCTGTCCCGCGCGCTGGTGGGC 375  
QY 127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146  
Db 376 GAGCAATGCGACACTGCGGGGCGCGCTTTCAGACTAACTGGATCTTCTGGGTGTGTGACA 435

QY 147 AspGlyProGlyAsnTyrLysTyrLysThrLysCysThrTrpLeuileGluGlyGlnPro 166  
Db 436 GATGACCTGGAAATTTAAATACAAAACGAAGTCACGTGGCTCATTTGAAGGACAGCA 495  
QY 167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186  
Db 496 AATAGATATATGAACTTCGTTTCAATCATTTTCTACAGAGTGTAGTTGGACATTTA 555  
QY 187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuile 206  
Db 556 TATGTTTATGATGGGACTCAATTTATGCACCGCTAGTTGTCTGATTTAGTGGCTCAT 615  
QY 207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226  
Db 616 GTTCTTGAGAGATGGCAATGAGACTGTCCCTGAGGTGTTCACCATCAGGTATATGCC 675  
QY 227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246  
Db 676 TTGCTGATTTTTTTAGTGATGCTGTATATTTGACTGGATTTAATATTACTTACAGT 735  
QY 247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266  
Db 736 TTTGATATGTCTCCAAATAACTGCTCAGGCGGAGAGAGTGTAAAGATCAGTAATAGCAGC 795  
QY 267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286  
Db 796 GA-TACTGTTGAATGTGAATGTTCTGAAAACCTGGAAGGTGAACATGTGACATTCCTCA 854  
QY 286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgG 306  
Db 855 CTGTACAGACAACATGTGGTTTTTCTCATCGAGGATCTGCAATTCAGATGATGTCAGAGG 914  
QY 306 YCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnG 326  
Db 915 ATGCTCTGCTTCTCAGACTGGCAGGGTCTGGATGTTTCAGTCTCTGTACAGCTAACCA 974  
QY 326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346  
Db 975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGAGCATCTCATAAGC 1034  
QY 346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366  
Db 1035 TGTGTCAATGGAACATTTATGTGGTGTGTGGAGGATATATGTTCAACCATCAGATTA 1094  
QY 366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386  
Db 1095 TAACATGTTCTAGCTATGACCTTGTCTTAGGAGTGGCTTCCATTAACCTTCTGT 1154  
QY 386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406  
Db 1155 GAACAATGTGGTGTGTAGATATGTCATTTCTTGGCATTTATACAGGATATAAATTTACAT 1214  
QY 406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426  
Db 1215 GTATGGAGGAAAAATTTGATTC-AACTGGGAATGACCAATGAGTTGAGAGTTTTCACA 1273  
QY 426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446  
Db 1274 TTCATAATGAGTCATGGTGTGTGTGACCCCTAAGGCAAGGAGCAGTATGACAGTGTG 1333  
QY 446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466  
Db 1334 GGCACTCTGCACACATTTGTACACTGAAGATGGCCAGTGGTGTGTCATCTGCTGTCATCTT 1393  
QY 466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486  
Db 1394 GTCACTGCCCTCTCTATGATATATAAGCAATGTGAGGAATATGATTTGGATAAGAACA 1453  
QY 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerV 506  
Db 1454 CATGGAGTATATTACACACCCAGGGTCCCTTGTGGAAGGGGGGTACGGCCATAGCAGTG 1513  
QY 506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaA 526



Db 1514 TTTACGACCATAGGACAGGCGCTATACGTTTCATGGTGGCTACAGGCTTTCAGTGCCA 1573  
QY 526 snlystYrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546  
Db 1574 ATAAGTACCGGCTTGCAGATGATCTCTACCGATATGATGCGATACCCAGATGTGGACCA 1633  
QY 546 leuLeuYsAspSerArgPheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566  
Db 1634 TTTCTTAAGGACACCGGATTTTCGGTTACTTGCACACAGCTGTGATAGTGAAGTGAACCA 1693  
QY 566 etLeuValPheGlyGlyValAsnThrHisAsnAspThrSerMetSerHisGlyAlaLeuYsCysP 586  
Db 1694 TGTGTGTGTGTGGGGGAACACACACATGACACATCTATGAGCCATGCGCCCAATGCT 1753  
QY 586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606  
Db 1754 TCTCTTTCAGATTTTCATGGCTATGACATTTGCCCTGTGACCGCTGGTCAGTGTTCGCCAGAC 1813  
QY 606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626  
Db 1814 CTGATCTCCACCAATGATGACACAGATTTGGCCATTCACAGCTCTTACACACACACCA 1873  
QY 626 etTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerG 646  
Db 1874 TGTATGTGTGTGTGTGTTCATAGTCTCTCTCCTCAGCGACATCTCTGTGTATTCACCTCGG 1933  
QY 646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
Db 1934 AACAGTGTGATGTCGCAATCGGAGTGAAGCCGCTGTGTTAGCAGAGACCTGTGTATTCGGT 1993  
QY 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686  
Db 1994 GTGTGTGGAACACACAGGCTGTCTCAGTGTATCTCGTGGCGCTGGCACTGATGAACAG 2053  
QY 686 luGluLeuYsLeuYsSerGluCysPheSerIysArgThrLeuAspHisAspArgCysAspG 706  
Db 2054 AAGAAAGATTTAAATCAGATGTTTTTCCAAAAGAACTCTTGACCATGACATGTGACC 2113  
QY 706 lnhIsthrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726  
Db 2114 AGACACAGATTTGACAGCTGACAGCCACACCAATGACTGCCACTGTGCAATGACC 2173  
QY 726 isCysValProArgAsnHisSerCysSerGluGlnIleSerIlePheArgTyrGluA 746  
Db 2174 ATTGTGTCCCGACGACACAGCTGTCTCAGAGGCCAGATCTCCATTTTATGATGAGA 2233  
QY 746 snCysProYsAspAsnProMetTyrTyrCysAsnYsIysThrSerCysArgSerCysA 766  
Db 2234 ATTGCCCCAAGGATAACCCCTATGTTACTGTGTAAACAGAGACACAGCTGCAGGAGCTGTG 2293  
QY 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuPro 784  
Db 2294 CCCTGGACCAAGAACTGCCAGTGGAGCCCGGAATCAGAGTGTGATGCTGCTGCC 2349

## RESULT 5

US-09-893-238-1  
; Sequence 1, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20

; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8827  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-1  
Alignment Scores:  
Pred. No.: 1,928-193 Length: 8827  
Score: 206.00 Matches: 576  
Percent Similarity: 96.00% Conservatives: 0  
Best Local Similarity: 96.00% Mismatches: 12  
Query Match: 14.42% Indels: 24  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-893-238-1 (1-8827)

QY 842 MetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLys 861  
Db 2790 ATGCAATCATCTCAAGTATGTCCAGCTCACTCTCACTCCATGGTGTGGTCTTCGGAAG 2849  
QY 862 IleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGln 881  
Db 2850 ATCAATGTGTCTTACTGTGTCTGGAGATATGTCTCATTCACAAATAGTTTGTGTCAG 2909  
QY 882 TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901  
Db 2910 TGGATGCCATCTGAGCCAGTGTGCTGTGGTCTCTGTGGATCTTGTTCAGAGCCTAGTACT 2969  
QY 902 ArgGlyLeuYsAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
Db 2970 CGGGGATTAAGGCTGCAACCTGCATCACTCTCAATGGCAGCGTCTGTGAAGCCCT 3029  
QY 922 AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941  
Db 3030 GCAACACACAGTGCACAGCAGTCCGACACCATGTGCTCCCTGCGGACAGCGTGTGGCCA 3088  
QY 942 CysThrSerGly-SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 961  
Db 3089 GTGCACCTAGCAG-CAGCTCGGAGTGCATGTGGTGCAGTAAACATGAACAGTGTGTGACT 3147  
QY 961 erAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTrpTyrThrMetSerT 981  
Db 3148 CCAATGCTACCTGCTCTCTCCCTTTGGCCAGTGTATGGATGGTATACGATGAGCA 3207  
QY 981 hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnPro 1000  
Db 3208 G-CTGCCCCACCTGAAAAATTGCTGTGGTACTGTACCTGCAGCCATTGCTTGGAGCAGCCA 3266  
QY 1001 GlyCysGlyTrpCysThrAspProSerAsnThrGlyLysGlyCysIleGluGlySer 1020  
Db 3267 GGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 3326  
QY 1021 TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039  
Db 3327 TATAAGGACCTGTGAAGATGCCCTCACAGGCCCTC-TG-CAGGAAATGT-GTATCCACAG 3383  
QY 1040 ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTrpSerPheIleHis 1059  
Db 3384 CCCCTTCTGAACTCCAGCATGTCTAGAGACAGCAGATCACTGGTCTTTCATTTCAC 3443  
QY 1060 CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079  
Db 3444 TGTCCAGCTTCCAGTGCACCGACACAGCAATGCATCAACAGAGATACTGTGAGAAG 3503  
QY 1080 CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099  
Db 3504 TGTGAGGA-CCTGACCCCGGCAAGCAGCTGCGAGACCTGTCATATCTGGCTTCTATGTGGA 3562  
QY 1099 pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119  
Db 3563 CCCGACTAATGGAGGCAAAATGTACGCCATGCAAGTGAATGGGACCGCATCACTGTGCAA 3622



QY 1119 nThrAsnThrGlyGlyCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnle 1139  
Db 3623 CACCACACCGGAGAGTCTTGTACCAACCAAGGTGCAAGGGGAGAGTGCAGCT 3682  
QY 1139 uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrThrL 1159  
Db 3683 ATGTGAGGTAGAAATCGATACCAAGMAACCTCTCAA-AGGAACATGCTACTATACCC 3741  
QY 1159 euLeuLeuAspTyrGlnPheThrPheSerLeuSerGlnLysAspArgTyrThrA 1179  
Db 3742 TTCTCATTCACATCATGTTACCTTTAGCTGTCCAGGAAGACACCGCTACTACAG 3801  
QY 1179 laileAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheleAsnA 1199  
Db 3802 CCATCAACTTTGTGGCTACTCTGATGACAAACACAGGGATTGGACATGTTTCATCATG 3861  
QY 1199 laSerLysAsnPhenLeuAnileThrTrpAlaAla-SerPheSer-AlaGlyThrGl 1218  
Db 3862 CTTCCAAAACCTTCAACCTCAACATCACCTGGGCCAC-CAGCTTCCC-AGCCGGAACCCA 3919  
QY 1218 nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237  
Db 3920 GAC-TGGAGAGAGGT-GCCTGTGTTCATAAACCAACATCAAGATACAAAGATAGC 3977  
QY 1238 PheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257  
Db 3978 TTCTCTAATGAGAAATTTGATTTTCGCAACCATCCAAACATCACATTTCTTTGTTATGTC 4037  
QY 1258 SerAsnPheThrTrpProLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277  
Db 4038 AGTAATTTTCACTTGGCCCATCAAAATTCAGATTGCTTCTCCAGCAGCAACCTTCATG 4097  
QY 1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAla 1297  
Db 4098 GACCTGGTACAGTCTTCTGACTTCTTCAGTTGTTCTCTCTGCTGCTGCTGCTGCT 4157  
QY 1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGluLeuLeu 1317  
Db 4158 CGAGTGGTCTGGAAGATCAAGCAGAGCTGTGGGCATCCAGCGGAGAGCAACTTCTT 4217  
QY 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337  
Db 4218 CGGAGATGCAACAGATGGCCAGCCGCCCTTTGCTTCTGTAAACGTTGCCCTTGGAAACA 4277  
QY 1338 AspGluGluProProAspLeuIleGlyGlySerIleLysThrValProLysProIleAla 1357  
Db 4278 GATGAGAGCTCTCTGATCTTATTTGGGGGAGTATAAAGACTGTTCCCAAAACCCATTGCA 4337  
QY 1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProArg 1377  
Db 4338 CTGGAGCCGTGTTTGGCAACAAAGCCGCTGCTCTCTGTTGTTGTGAGGCTCCCTCGA 4397  
QY 1378 GlyLeuGlyGlyLeuProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397  
Db 4398 GGCTGGGTGGCATCCCTCTCTGGGCAGTCAGGTCTTGCTGTGGCAGCCCTGCTGTG 4457  
QY 1398 AspIleSerGlnGlnMetProLysValTyrLysGluLysSerGlyAlaValArgAsnArg 1417  
Db 4458 GACATTTCTCAGCAGATGCCGATGTGTACAGGAGAACTCAGGAGCCGTGTGAAACCCGG 4517  
QY 1418 LysGlnGlnProAlaGlnProGlyThrCysIle 1429  
Db 4518 AAGCAGACGCCCCCTGACAGCTGGGACCTGCATC 4553

## RESULT 6

US-09-893-238-8  
; Sequence 8, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND

; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-8  
Alignment Scores:  
Pred. No.: 9,14e-83 Length: 2419  
Score: 94.00 Matches: 94  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.58% Indels: 0  
DB: Gaps: 0  
US-09-787-097-12 (1-1429) x US-09-893-238-8 (1-2419)  
QY 513 AlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAsp 532  
Db 1803 GCTCTGACGTTCATGTGGCTTACAGGCTTACGCCCAACAAATACCGGCTTGCAGAT 1862  
QY 533 AspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhe 552  
Db 1863 GACCTTACAGATACGATGTCGATCTCAGATGTCGACCATTTCTTAAGGACAGCCGATTT 1922  
QY 553 PheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyGlyAsn 572  
Db 1923 TTCGTTACTTGCATACAGCTGTGATAGTAGTGAGGACCATGCTGTGTTGGAGGGAAC 1982  
QY 573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerAspPheMetAla 592  
Db 1983 ACACACATGACACTTCCATGACGACGTCGACCATGCTTCTCCTCGGACTTCATGGCT 2042  
QY 593 TyrAspIleAlaCysAspArgTyrTrpSerValLeuProArgPro 606  
Db 2043 TATGACATTGCTGTGACCGATGTCAGTGTGCTTCCCGACACCT 2084  
RESULT 7  
US-10-198-846-5869/c  
; Sequence 5869, Application US/10198846  
; Publication No. US20030099974A1  
; GENERAL INFORMATION:  
; APPLICANT: Lillie, James  
; APPLICANT: Xu, Yongyao  
; APPLICANT: Wang, Youzhen  
; APPLICANT: Steimann, Kathleen  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS  
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF BREAST CANCER  
; FILE REFERENCE: MRI-049  
; CURRENT APPLICATION NUMBER: US/10/198,846  
; CURRENT FILING DATE: 2002-07-18  
; PRIOR APPLICATION NUMBER: 60/306,220  
; PRIOR FILING DATE: 2001-07-18  
; NUMBER OF SEQ ID NOS: 14084  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5869  
; LENGTH: 625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature

; LOCATION: 3, 4, 5, 6, 7, 8, 10, 12, 14, 15, 17, 26, 28, 29  
; OTHER INFORMATION: n = A,T,C or G  
US-10-198-846-5869

Alignment Scores:  
Pred. No.: 2,48e-81 Length: 625  
Score: 92.00 Matches: 92  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.44% Indels: 0  
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-198-846-5869 (1-625)

Qy 901 ThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArg 920  
Db 338 ACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGGTAGTCTGTGAAAGG 279  
Qy 921 ProAlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGly 940  
Db 278 CTGCGAAACCACTGCTCTAAGGCTGCCGACACCATGTCCTTGAGGACAGCATGTGGA 219  
Qy 941 AspCysThrSerGlySerGluCysMetTrpCysSerAsnMetLysGlnCysValAsp 960  
Db 218 GATTGACCAGCGGACGCTGAGTGTATGCTGTCGAGCAACATGAAGCAGTGTGTGGAC 159  
Qy 961 SerAsnAlaTrpValAlaSerPheProPheGlyGlnCysMetGluTrpTrpThrMetSer 980  
Db 158 TCCAAATGCCATGTGGCTTCCTTCCTTTGGCAGTGTATGATGGATGTATGATGAGC 99  
Qy 981 ThrCysProGluAsnCysSerGlyTrpCysThr 992  
Db 98 ACCTGCCCCCTGAAATTTGTTGAGGCTACTGTACC 63

## RESULT 8

US-09-736-457-966/c  
; Sequence 966, Application US/09736457  
; Patent No. US20020168637A1

GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong

; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.

; APPLICANT: Fanger, Gary

; APPLICANT: Vedvick, Tom

; APPLICANT: Carter, Darrick

; APPLICANT: Retter, Marc

; APPLICANT: Mannion, Jane

; APPLICANT: Fan, Liqun

; APPLICANT: Wang, Aijun

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C15

CURRENT APPLICATION NUMBER: US/09/736,457

CURRENT FILING DATE: 2000-12-13

NUMBER OF SEQ ID NOS: 1864

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 966

LENGTH: 246

TYPE: DNA

ORGANISM: Homo sapien

US-09-736-457-966

Alignment Scores:  
Pred. No.: 8,74e-72 Length: 246  
Score: 82.00 Matches: 82  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.74% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-736-457-966 (1-246)

Qy 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858

Db 246 CTGCGAATATGAGTCATCTCAGACATGTCCTCAAGCTCACCTTAACCCCATGGTGGC 187  
Qy 859 LeuArgLysIleAsnValSerTrpCysTrpGluAspMetSerProPheThrAsnSer 878  
Db 186 CTTGCGAAGATCAATGTGCTCTACTGCTGGTGGGAAGATATGTCCTCAATTTACAAATAGT 127  
Qy 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898  
Db 126 TTACTACAGTGGATCCGCTCTGAGCCCGAGTGAATCTGTGGATTTTATCAGAA 67  
Qy 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918  
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGAGTAGTGTCTGT 7

Qy 919 GluArg 920

Db 6 GAAAGG 1

## RESULT 9

US-09-902-941-966/c

; Sequence 966, Application US/09902941

; Patent No. US20020172952A1

GENERAL INFORMATION:

; APPLICANT: Henderson, Robert A.

; APPLICANT: Wang, Tongtong

; APPLICANT: Watanabe, Yoshihiro

; APPLICANT: Johnson, Jeffrey C.

; APPLICANT: Retter, Marc W.

; APPLICANT: Warnerakis, Margarita

; APPLICANT: Carter, Darrick

; APPLICANT: Fanger, Gary R.

; APPLICANT: Vedvick, Thomas S.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: McNabb, Andria

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

AND DIAGNOSIS OF LUNG CANCER

FILE REFERENCE: 210121.478C17

CURRENT APPLICATION NUMBER: US/09/902,941

CURRENT FILING DATE: 2001-07-10

NUMBER OF SEQ ID NOS: 2002

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 966

LENGTH: 246

TYPE: DNA

ORGANISM: Homo sapiens

US-09-902-941-966

## Alignment Scores:

Pred. No.: 8,74e-72 Length: 246  
Score: 82.00 Matches: 82  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.74% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-902-941-966 (1-246)

Qy 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858  
Db 246 CTGCGAATATGAGTCATCTCAGACATGTCCTCAAGCTCACCTTAACCCCATGGTGGC 187  
Qy 859 LeuArgLysIleAsnValSerTrpCysTrpGluAspMetSerProPheThrAsnSer 878  
Db 186 CTTGCGAAGATCAATGTGCTCTACTGCTGGTGGGAAGATATGTCCTCAATTTACAAATAGT 127  
Qy 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898  
Db 126 TTACTACAGTGGATCCGCTCTGAGCCCGAGTGAATCTGTGGATTTTATCAGAA 67  
Qy 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918  
Db 66 CCCAGTACTCGGGGACTGAAGGCTGCAACCTGCATCAACCCACTCAATGAGTAGTGTCTGT 7

246

Score: 82.00 Matches: 82  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.74% Indels: 0  
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-113-872-966 (1-246)

QY 839 LeuArgIleMetGlnSerGlnSerMetSerLysLeuThrLeuThrProTrpValGly 858  
DB 246 CTGCGAATAATGAGTCATCTCAGAGCATGTCAGAGCTCACCTTAACCCATGGGTGGC 187  
QY 859 LeuArgLysIleAsnValSerTrpCysTrpGluAspMetSerProPheThrAsnSer 878  
DB 186 CTTCGGAAGATCAATGTGCTACTGTCTGGAGATATGCTCCCATTTACAATAGT 127  
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898  
DB 126 TTACTACAGTGGATGCGTCTGAGCCAGCTGATGCTGGATTTCTGGAAATTTATCAGAA 67  
QY 899 ProSerThrArgGlyLeuLysAlaIleThrCysIleAsnProLeuAsnGlySerValCys 918  
DB 66 CCAGTACTCGGGAGCTGAAGGCTGCACTGCATCAACCCCATCAATGGTAGTGTCTGT 7  
QY 919 GluArg 920  
DB 6 GAAAGG 1

## RESULT 13

US-10-277-216-5  
; Sequence 5, Application US/10277216  
; Publication No. US2004002470A1  
; GENERAL INFORMATION:  
; APPLICANT: KEITH, TIM  
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,  
; FILE REFERENCE: 2976-4051  
; CURRENT APPLICATION NUMBER: US/10/277,216  
; PRIOR FILING DATE: 2002-10-17  
; PRIOR APPLICATION NUMBER: 10/126,022  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR FILING DATE: 2002-04-19  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR FILING DATE: 2000-04-13  
; NUMBER OF SEQ ID NOS: 420  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 207433  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-277-216-5

Alignment Scores:  
Pred. No.: 4.58e-64 Length: 207433  
Score: 77.00 Matches: 77  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.39% Indels: 0  
DB: 15 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-277-216-5 (1-207433)

QY 985 GluAsnCysSerGlyTrpCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004  
DB 9815 GAAATTTGTTTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGG 9874  
QY 1005 CysThrAspProSerAsnThrGlyLysGlyCysIleGluGlySerTrpLysGlyPro 1024  
DB 9875 TGTAATGATCCAGCAATCTGCAAGGAAATGATAGAGGTTCTCTTAAGGACCA 9934  
QY 1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTrpGlnProLeuLeuAsnSer 1044

DB 9935 GTGAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGTCTCAATTCC 9994  
QY 1045 SerMetCysLeuGluAspSerArgTrpAsnTrpSerPheIleHisCysPro 1061  
DB 9995 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCATTCACTGTCCA 10045

## RESULT 14

US-10-126-022-5  
; Sequence 5, Application US/10126022  
; Publication No. US20040023215A1  
; GENERAL INFORMATION:  
; APPLICANT: KEITH, TIM  
; TITLE OF INVENTION: NOVEL HUMAN GENE RELATING TO RESPIRATORY DISEASES,  
; FILE REFERENCE: 2976-4039US2  
; CURRENT APPLICATION NUMBER: US/10/126,022  
; CURRENT FILING DATE: 2002-04-19  
; PRIOR APPLICATION NUMBER: 09/834,597  
; PRIOR FILING DATE: 2001-04-13  
; PRIOR APPLICATION NUMBER: 09/548,797  
; PRIOR FILING DATE: 2000-04-13  
; NUMBER OF SEQ ID NOS: 420  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 207433  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-126-022-5

## Alignment Scores:

Pred. No.: 4.58e-64 Length: 207433  
Score: 77.00 Matches: 77  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.39% Indels: 0  
DB: 16 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-126-022-5 (1-207433)

QY 985 GluAsnCysSerGlyTrpCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrp 1004  
DB 9815 GAAATTTGTTTCAGGCTACTGTACTGTAGTCAATGCTTGGAGCAACAGGCTGTGGCTGG 9874  
QY 1005 CysThrAspProSerAsnThrGlyLysGlyCysIleGluGlySerTrpLysGlyPro 1024  
DB 9875 TGTAATGATCCAGCAATCTGCAAGGAAATGATAGAGGTTCTCTTAAGGACCA 9934  
QY 1025 ValLysMetProSerGlnAlaProThrGlyAsnPheTrpGlnProLeuLeuAsnSer 1044  
DB 9935 GTGAGATGCTTTCGCAAGCCCTACAGGAAATTTCTATCCACAGCCCTGTCTCAATTCC 9994  
QY 1045 SerMetCysLeuGluAspSerArgTrpAsnTrpSerPheIleHisCysPro 1061  
DB 9995 AGCATGTGTCTAGAGGACAGCAGATACAACTGGTCTTTTCATTCACTGTCCA 10045

## RESULT 15

US-09-893-238-10  
; Sequence 10, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20

```
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-10

Alignment Scores:
Pred. No.:      3,116-65      Length:      1051
Score:          76.00      Matches:      111
Percent Similarity: 98.23%      Conservative: 0
Best Local Similarity: 98.23%      Mismatches: 1
Query Match:      5.32%      Indels:      2
DB:              9          Gaps:      0

US-09-787-097-12 (1-1429) x US-09-893-238-10 (1-1051)

Qy 87 AlaAlaValSerGlySerAlaAlaGluAlaLysGluCysAspArgProCysVal 106
Db 463 GCGCGCGGGTGTCCGGCTCGGCCGAGCCAGAGGCAAGGAATGTGACCGCGCGTGTGC 522

Qy 107 AsnGlyGlyArgCysAsnProGlyThrGlyGlnCysValCys-ProAlaGlyTrpValG1 126
Db 523 AACGCGCGCGGTGTGCAACCTCGCCACCGCCAGTGGTGTGCCACG-GGCTGGGTGGG 581

Qy 126 yGluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerSerGlyPheValTh 146
Db 582 CGAGCAATGCCAGCACTCGGGGGCGCTTCAGACTAACTGGCTCTCTGGATTGTAAAC 641

Qy 146 rAspGlyProGlyAsnThrLysThrLysThrLysCysThrTrpLeuLleGluGlyGlnPr 166
Db 642 AGATGGACCTGGGAATTAAATAAAGACGAAGTGCACATGGCTTCATTGAAGGACACCC 701

Qy 166 oAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLe 186
Db 702 AAATAGATATATGAGACTTCGCTTCAACCATTTTCTACAGATGTAGCTGGACCATTT 761

Qy 186 uTrpValTrpAspGlyAspSerIleTyrAlaProLeu 198
Db 762 ATATGTTTATGATGGGACTCAATCTAGCACCTCTG 798

RESULT 16
US-09-893-238-5
; Sequence 5, Application US/09893238
; Patent No. US20020150973A1
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND
; FILE REFERENCE: 7853-237
; CURRENT APPLICATION NUMBER: US/09/893,238
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: 09/245,041
; PRIOR FILING DATE: 1999-02-05
; PRIOR APPLICATION NUMBER: 60/093,630
; PRIOR FILING DATE: 1998-07-21
; PRIOR APPLICATION NUMBER: 60/104,978
; PRIOR FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 129
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-893-238-5

Alignment Scores:
Pred. No.:      1,966-60      Length:      90050
Score:          73.00      Matches:      73
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
```

```
Query Match:      5.11%      Indels:      0
DB:              9          Gaps:      0

US-09-787-097-12 (1-1429) x US-09-893-238-5 (1-90050)

Qy 849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerTrpCys 868
Db 28996 TCCAAGCTCACTCTGACTCGGTGGTTGGTCTTCGGAAGATCAATGTGTCTTACTGTGC 29055

Qy 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
Db 29056 TGGAGGATATGTCTTCATTCACAAATAGTTTCTCAGTGGATGCCATCTGAGCCAGT 29115

Qy 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
Db 29116 GATGCTGGCTTCTGTGGATCTTGTTCAGAGCCTAGTACTCGGGATTAAGGCTGCAACC 29175

Qy 909 CysIleAsnProLeuAsnGlySerValCysGluArgPro 921
Db 29176 TGCATCAACCTCTCAATGGCAGCGTCTGTGAAGGCGCT 29214

RESULT 17
US-09-864-761-15900
; Sequence 15900, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,697
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
```

; SEQ ID NO 15900  
; LENGTH: 531  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL132773.14  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
US-09-864-761-15900

Alignment Scores:  
Pred. No.: 1,466-58 Length: 531  
Score: 69.00 Matches: 69  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.83% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-864-761-15900 (1-531)

QY 414 ThrGlyAsnValThrAsnGluLeuValPheHisIleHisAsnGluSerTrpValLeu 433  
Db 186 ACTGGGAATGTGACCAATGAGTGGAGTTTTCACATTCAATGAGTCATGGGTGGT 245  
QY 434 LeuThrProLysAlalysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThr 453  
Db 246 TTGACCCCTAAGCAAGGAGCAGTATGCAATGAGTGGTGGGCACCTCTGCACACATTGTTACA 305  
QY 454 LeuLysAsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyr 473  
Db 306 CTGAAGAATGGCGGAGTGGTGCATCTTGGTCACTCTTGGTCACTGGCCCTCTCTATGGATAT 365  
QY 474 IleSerAsnValGlnGluTyrAspLeu 482  
Db 366 ATAAGCAATGTGCAGGAATATGATTG 392

## RESULT 18

US-09-864-761-32404  
; Sequence 32404, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663

; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 32404  
; LENGTH: 193  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AL132773.14  
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.9  
; OTHER INFORMATION: EST HUMAN HIT: BF313142.1, EVALUE 1.00e-104  
; OTHER INFORMATION: NT HIT: AF218896.1, EVALUE 1.00e-105  
; OTHER INFORMATION: SWISSPROT HIT: Q19981, EVALUE 1.00e-05  
US-09-864-761-32404

Alignment Scores:  
Pred. No.: 5,19e-54 Length: 193  
Score: 64.00 Matches: 64  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 4.48% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-864-761-32404 (1-193)

QY 416 AsnValThrAsnGluLeuArgValPheHisIleHisAsnGluSerTrpValLeuLeuThr 435  
Db 2 AATGTGACCAATGAGTGGAGTTTTCACATTCAATGAGTCATGGGTGGTGGACC 61  
QY 436 ProLysAlalysGluGlnTyrAlaValValGlyHisSerAlaHisIleValThrLeuLys 455  
Db 62 CCTAAGCAAGGAGCAGTATGCAATGAGTGGTGGGCACCTCTGCACACATTGTACACTGAAG 121  
QY 456 AsnGlyArgValValMetLeuValIlePheGlyHisCysProLeuTyrGlyTyrIleSer 475  
Db 122 AATGGCCGAGTGGTCATGCTGGTCACTTTGGTCACTGGCCCTCTCTATGGATATATAAGC 181  
QY 476 AsnValGlnGlu 479  
Db 182 AATGTGCAGGAA 193

## RESULT 19

US-09-864-761-21460/c  
; Sequence 21460, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US/09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366

```
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 21460
/ LENGTH: 182
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO A132773.1
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
/ OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
/ OTHER INFORMATION: NT HIT: AF218902.1, EVALUE 2.00e-98
/ OTHER INFORMATION: EST HUMAN HIT: A1372810.1, EVALUE 2.00e-98
/ OTHER INFORMATION: SWISSPROT HIT: P22897, EVALUE 8.00e-03
US-09-864-761-21460
```

```
Alignment Scores:
Pred. No.: 4,6e-50 Length: 182
Score: 60.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 4.20% Indels: 0
DB: 9 Gaps: 0
```

US-09-787-097-12 (1-1429) x US-09-864-761-21460 (1-182)

```
Qy 849 SerIysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValserTyrTrpCys 868
Db 182 TCCAGCTCACCTTAACCCCATGGTCCGCTCGGAAGATCAATGTGTCTACTGGTGC 123
Qy 869 TrpGluaspMetSerProthethraanSerLeuLeuGlnTrpMetProSerGluProSer 888
Db 122 TGGGAAGATGTGCCCATTTACAAATAGTTTACACAGTGGATCCGCTCAGGCCAGT 63
Qy 889 AspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaIathr 908
```

Db 62 GATGCTGGATTCTGTGGATTATCAGAACCCAGTACTCGGCGGACTGAGGCTGCAACC 3

# RESULT 20

```
US-09-864-761-4720/c
/ Sequence 4720, Application US/09864761
/ Patent No. US20020048763A1
/ GENERAL INFORMATION:
/ APPLICANT: Penn, Sharon G.
/ APPLICANT: Rank, David R.
/ APPLICANT: Hanzel, David K.
/ APPLICANT: Chen, Wensheng
/ TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
/ FILE REFERENCE: Aeomica-X-1
/ CURRENT APPLICATION NUMBER: US/09/864,761
/ CURRENT FILING DATE: 2001-05-23
/ PRIOR APPLICATION NUMBER: US 60/180,312
/ PRIOR FILING DATE: 2000-02-04
/ PRIOR APPLICATION NUMBER: US 60/207,456
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: US 09/632,366
/ PRIOR FILING DATE: 2000-08-03
/ PRIOR APPLICATION NUMBER: GB 24263.6
/ PRIOR FILING DATE: 2000-10-04
/ PRIOR APPLICATION NUMBER: US 60/236,359
/ PRIOR FILING DATE: 2000-09-27
/ PRIOR APPLICATION NUMBER: PCT/US01/00666
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00667
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00664
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00669
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00665
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00668
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00663
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00662
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00661
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: PCT/US01/00670
/ PRIOR FILING DATE: 2001-01-30
/ PRIOR APPLICATION NUMBER: US 60/234,687
/ PRIOR FILING DATE: 2000-09-21
/ PRIOR APPLICATION NUMBER: US 09/608,408
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: US 09/774,203
/ PRIOR FILING DATE: 2001-01-29
/ NUMBER OF SEQ ID NOS: 49117
/ SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
/ SEQ ID NO 4720
/ LENGTH: 375
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: MAP TO A132773.1
/ OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.96
/ OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1
/ OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.4
/ OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.6
/ OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
/ OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.2
/ OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
/ OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1
US-09-864-761-4720
```



Alignment Scores: 8.65e-47 Length: 375  
Pred. No.: 57.00 Matches: 57  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 3.99% Gaps: 0  
DB:

US-09-787-097-12 (1-1429) x US-09-864-761-4720 (1-375)

QY 865 SerTyrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetPro 884  
Db 374 TCTACTGTGCTGGGAAGATATGTCCTCATTTACAAATAGTTTACTACAGTGGATCGCG 315  
QY 885 SerGluProSerAspIleGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeu 904  
Db 314 TCTAGCCCGATGATCTGATCTGTGGAATTTATCAGAACCCAGTACTCGGGGACTG 255  
QY 905 LysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921  
Db 254 AGGCTGCAACCTGCATCAACCACCTCATGTAGTGTCTGTGAAGGCT 204

RESULT 21

US-09-893-238-6/c  
; Sequence 6, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/945,041  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-10-20  
; PRIOR APPLICATION NUMBER: 60/104,978  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 6  
; LENGTH: 226  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-6

Alignment Scores: 4.87e-41 Length: 226  
Pred. No.: 51.00 Matches: 51  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 3.57% Gaps: 0  
DB:

US-09-787-097-12 (1-1429) x US-09-893-238-6 (1-226)

QY 871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890  
Db 225 GATATGTCCTCATTCACAAATAGTTTGTGAGTGGATGCCATCTGAGCCAGTGTGCT 166  
QY 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLeuAlaAlaThrCysIle 910  
Db 165 GGTCTTCTGGGATCTGTGACAGCCCTAGTATCTCGGGGATTAAAGGCTGCAACCTGCATC 106  
QY 911 AsnProLeuAsnGlySerValCysGluArgPro 921  
Db 105 AACCTCTCAATGGCAGCGCTCTGTGAAGGCT 73

RESULT 22

US-09-796-692-7610  
; Sequence 7610, Application US/09796692

Publication No. US20020198362A1  
; GENERAL INFORMATION:  
; APPLICANT: Gauger, Alexander  
; APPLICANT: Algate, Paul A.  
; APPLICANT: Mannion, Jane  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE DETECTION, DIAGNOSIS AND THERAPY  
; FILE REFERENCE: 2077.001200  
; CURRENT APPLICATION NUMBER: US/09/796,692  
; CURRENT FILING DATE: 2001-03-01  
; PRIOR APPLICATION NUMBER: 60/186,126  
; PRIOR FILING DATE: 2000-03-01  
; PRIOR APPLICATION NUMBER: 60/190,479  
; PRIOR FILING DATE: 2000-03-17  
; PRIOR APPLICATION NUMBER: 60/200,545  
; PRIOR FILING DATE: 2000-04-27  
; PRIOR APPLICATION NUMBER: 60/200,303  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,779  
; PRIOR FILING DATE: 2000-04-28  
; PRIOR APPLICATION NUMBER: 60/200,999  
; PRIOR FILING DATE: 2000-05-01  
; PRIOR APPLICATION NUMBER: 60/202,084  
; PRIOR FILING DATE: 2000-05-04  
; PRIOR APPLICATION NUMBER: 60/206,201  
; PRIOR FILING DATE: 2000-05-22  
; PRIOR APPLICATION NUMBER: 60/218,950  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/222,903  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: 60/223,416  
; PRIOR FILING DATE: 2000-08-04  
; PRIOR APPLICATION NUMBER: 60/223,378  
; PRIOR FILING DATE: 2000-08-07  
; NUMBER OF SEQ ID NOS: 9597  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7610  
; LENGTH: 212  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (140)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (142)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (162)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (196)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (204)  
; OTHER INFORMATION: n=A,T,C or G  
; NAME/KEY: unsure  
; LOCATION: (205)  
; OTHER INFORMATION: n=A,T,C or G  
US-09-796-692-7610  
Alignment Scores: 4.23e-36 Length: 212  
Pred. No.: 46.00 Matches: 46  
Score: 100.00% Conservative: 0  
Percent Similarity: 100.00% Mismatches: 0  
Best Local Similarity: 100.00% Indels: 0  
Query Match: 3.22% Gaps: 0  
DB: 9  
US-09-787-097-12 (1-1429) x US-09-796-692-7610 (1-212)

QY 972 GlnCysMetGluTrpTyrThrMetSerThrCysProGluAsnCysSerGlyTyrCys 991

```
Db 2 CAGTGTATGTAAGTATACGATGACACCTGCCCTCCCTGAAATTTGTTCAAGCTACTGT 61
Qy 992 ThrCysSerHisCysLeuGluInProGlyCysGlyTrpCysThrAspProSerAsnThr 1011
Db 62 ACCTGTAGTCATTGCTTGGAGCAACAGGCTGTGGTGTGTACTGATCCAGCAATACT 121
Qy 1012 GlyLysGlyLysCysalle 1017
Db 122 GGCAGGAGGAATGCATA 139

RESULT 23
US-10-040-862-7610
; Sequence 7610, Application US/10040862
; Publication No. US20030078396A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-013520US
; CURRENT APPLICATION NUMBER: US/10/040,862
; CURRENT FILING DATE: 2001-11-06
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: US 60/223,416
; PRIOR FILING DATE: 2000-08-04
; PRIOR APPLICATION NUMBER: US 60/223,378
; PRIOR FILING DATE: 2000-08-07
; PRIOR APPLICATION NUMBER: US 09/796,692
; PRIOR FILING DATE: 2001-03-01
; NUMBER OF SEQ ID NOS: 10467
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7610
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (140)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (142)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (162)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (196)
; OTHER INFORMATION: n=A,T,C or G
```

```
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (204)
; OTHER INFORMATION: n=A,T,C or G
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (205)
; OTHER INFORMATION: n=A,T,C or G
US-10-040-862-7610

Alignment Scores:
Pred. No.: 4.23e-36 Length: 212
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.22% Indels: 0
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-040-862-7610 (1-212)
Qy 972 GlnCysMetGluTptYrThrMetSerThrCysProGluAsnCysSerGlyTrpCys 991
Db 2 CAGTGTATGTAAGTATACGATGACACCTGCCCTCCCTGAAATTTGTTCAAGCTACTGT 61
Qy 992 ThrCysSerHisCysLeuGluInProGlyCysGlyTrpCysThrAspProSerAsnThr 1011
Db 62 ACCTGTAGTCATTGCTTGGAGCAACAGGCTGTGGTGTGTACTGATCCAGCAATACT 121
Qy 1012 GlyLysGlyLysCysalle 1017
Db 122 GGCAGGAGGAATGCATA 139

RESULT 24
US-10-057-475B-7610
; Sequence 7610, Application US/10057475B
; Publication No. US20040002068A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Clapper, Jonathan David
; APPLICANT: Wang, Aijun
; APPLICANT: Ordenez, Nadia
; APPLICANT: Carter, Lauren
; APPLICANT: McNeill, Patricia Dianne
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; TITLE OF INVENTION: Hematological Malignancies
; FILE REFERENCE: 014058-014402US
; CURRENT APPLICATION NUMBER: US/10/057,475B
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 10979
```

```
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7610
; LENGTH: 212
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(212)
; OTHER INFORMATION: n = g, a, c or t
US-10-057-475B-7610

Alignment Scores:
Pred. No.:      4,23e-36      Length:      212
Score:          46.00        Matches:      46
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     3.22%       Indels:      0
DB:              15         Gaps:       0

US-09-787-097-12 (1-1429) x US-10-057-475B-7610 (1-212)
QY      972 GlnCysMetGluTrpTyrThrMetSerThrCysProGluAsnCysSerGlyTyrCys 991
      |||
Db      2 CAGTGTATGGAATGGTATACGATCAGCACCTGCCCTGAAAAATTGTTTCAGGCTACTGT 61
QY      992 ThrCysSerHisCysLeuGlnProGlyCysGlyTyrCysThrAspProSerAsnThr 1011
      |||
Db      62 ACCTGTAGTCATTGCTGGAGCAACAGGCTGGCTGGTGTACTGATCCAGCAATACT 121
QY      1012 GlyLySGlyLysCysIle 1017
      |||
Db      122 GGCAAGGGGAATGCATA 139

RESULT 25
US-10-154-884B-7610
; Sequence 7610, Application US/10154884B
; Publication No. US20040005581A1
; GENERAL INFORMATION:
; APPLICANT: Gaiger, Alexander
; APPLICANT: Algate, Paul A.
; APPLICANT: Mannion, Jane
; APPLICANT: Retter, Marc W.
; APPLICANT: Corixa Corporation
; TITLE OF INVENTION: Compositions and Methods for the Detection, Diagnosis and Therapy
; FILE REFERENCE: 014058-013521US
; CURRENT APPLICATION NUMBER: US/10/154,884B
; CURRENT FILING DATE: 2002-05-23
; PRIOR APPLICATION NUMBER: US 60/186,126
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: US 60/190,479
; PRIOR FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: US 60/200,545
; PRIOR FILING DATE: 2000-04-27
; PRIOR APPLICATION NUMBER: US 60/200,303
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,779
; PRIOR FILING DATE: 2000-04-28
; PRIOR APPLICATION NUMBER: US 60/200,999
; PRIOR FILING DATE: 2000-05-01
; PRIOR APPLICATION NUMBER: US 60/202,084
; PRIOR FILING DATE: 2000-05-04
; PRIOR APPLICATION NUMBER: US 60/206,201
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: US 60/218,950
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: US 60/222,903
; PRIOR FILING DATE: 2000-08-03
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 11290
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7610
; LENGTH: 212

; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(212)
; OTHER INFORMATION: n = g, a, c or t
US-10-154-884B-7610

Alignment Scores:
Pred. No.:      4,23e-36      Length:      212
Score:          46.00        Matches:      46
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match:     3.22%       Indels:      0
DB:              15         Gaps:       0

US-09-787-097-12 (1-1429) x US-10-154-884B-7610 (1-212)
QY      972 GlnCysMetGluTrpTyrThrMetSerThrCysProGluAsnCysSerGlyTyrCys 991
      |||
Db      2 CAGTGTATGGAATGGTATACGATCAGCACCTGCCCTGAAAAATTGTTTCAGGCTACTGT 61
QY      992 ThrCysSerHisCysLeuGlnProGlyCysGlyTyrCysThrAspProSerAsnThr 1011
      |||
Db      62 ACCTGTAGTCATTGCTGGAGCAACAGGCTGGCTGGTGTACTGATCCAGCAATACT 121
QY      1012 GlyLySGlyLysCysIle 1017
      |||
Db      122 GGCAAGGGGAATGCATA 139

RESULT 26
US-09-864-761-3681/c
; Sequence 3681, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aescmca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
```

```

; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
; SEQ ID NO 3681
; LENGTH: 402
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL132773.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
US-09-864-761-3681

Alignment Scores:
Pred. No.: 7,72e-36 Length: 402
Score: 46.00 Matches: 46
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.22% Indels: 0
DB: Gaps: 0

US-09-787-097-12 (1-1429) x US-09-864-761-3681 (1-402)

QY 1062 AlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGlu 1081
Db 389 GCTTGCCAAATGCAACGGCCACAGTAATCATCAATCATGAGCATCTGTGAGAAGTGTGAG 330
QY 1082 AsnLeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyGlyAspProThr 1101
Db 329 AACCTCACCACAGCGCAAGCACTGCGAGACCTGCATATCTGGCTTCTACGGTGATCCCCACC 270
QY 1102 AsnGlyGlyLysCysGln 1107
Db 269 AATGGAGGGAAATGTCTAG 252

RESULT 27
US-09-864-761-20449/c
; Sequence 20449, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Weisheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-x-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27

```

```

; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annapex Sequence Listing Engine vers. 1.1
; SEQ ID NO 20449
; LENGTH: 137
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO ALL32773.1
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.9
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.1
; OTHER INFORMATION: EST HUMAN HIT: BF672370.1, EVALUATE 2.00e-71
; OTHER INFORMATION: NT HIT: AB011120.1, EVALUATE 1.00e-71
; OTHER INFORMATION: SWISSPROT HIT: Q19981, EVALUATE 3.00e-11
US-09-864-761-20449

Alignment Scores:
Pred. NO.: 2.76e-35 Length: 137
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.15% Indels: 0
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-864-761-20449 (1-137)

Qy 1063 CysGlnCysAsnGlyHisSerLysCysFilaSnglnSerIleCysGluLysCysGluAsn 1082
Db 135 TGGCAATGTCACGGCCACAGTAATGATCATCATCATAGACATCTGTGAGACTGTGAGAAC 76
Qy 1083 LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsn 1102
Db 75 CTGACCACAGCCACAGCAGCTCGAGACCTGCATATCTGCCTTCTACGCTGATCCCAACAT 16
Qy 1103 GlyGlyLysCysGln 1107
Db 15 GGACGGAAATGTCTAG 1

RESULT 28
US-09-893-238-7/c

```

; Sequence 7, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 7  
; LENGTH: 221  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-7

Alignment Scores:  
Pred. No.: 4,32e-35 Length: 221  
Score: 45.00 Matches: 45  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.15% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-893-238-7 (1-221)

QY 871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890  
|||  
DB 220 GATATGCTCCATCAACATAATGTTGCTGCAGTGGATGCCATCTGAGCCAGTGTGCT 161  
|||

QY 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910  
|||  
DB 160 GGGTTCTGTGGGATCTGTGTCAGAGCCTAGTACTCGGGGATTAAAGCTGCAACCTGCATC 101  
|||

QY 911 AsnProLeuAsnGly 915  
|||  
DB 100 AACCTCTCAATGGT 86  
|||

RESULT 29  
US-09-893-238-4  
; Sequence 4, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 4  
; LENGTH: 5973  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-4

Alignment Scores:  
Pred. No.: 4,32e-35 Length: 221  
Score: 45.00 Matches: 45  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 3.15% Indels: 0  
DB: 9 Gaps: 0

Pred. No.: 7,92e-22 Length: 5973  
Score: 33.00 Matches: 33  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.31% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-893-238-4 (1-5973)

QY 166 ProAsnArgIleMetArgLeuAtcGpHeAsnHisPheAlaThrGluCysSerTrpAspHis 185  
|||  
DB 4328 CCAATAGATPATGAGACTTCGCTTCACCACTTTCTCAGATGTAAGTGGACCAT 4387  
|||

QY 186 LeuTrpValTyArgGlyAspSerIleTyAlaProLeu 198  
|||  
DB 4388 TTATATGTTTATGATGGGACTCAATCTACGCACCTCTG 4426  
|||

RESULT 30  
US-09-893-238-3/c  
; Sequence 3, Application US/09893238  
; Patent No. US20020150973A1  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE TREATMENT AND  
; TITLE OF INVENTION: DIAGNOSIS OF BODY WEIGHT DISORDERS, INCLUDING OBESITY  
; FILE REFERENCE: 7853-237  
; CURRENT APPLICATION NUMBER: US/09/893,238  
; CURRENT FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: 09/245,041  
; PRIOR FILING DATE: 1999-02-05  
; PRIOR APPLICATION NUMBER: 60/093,630  
; PRIOR FILING DATE: 1998-07-21  
; PRIOR APPLICATION NUMBER: 60/104,978  
; PRIOR FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 129  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 17056  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-893-238-3

Alignment Scores:  
Pred. No.: 2,03e-18 Length: 17056  
Score: 30.00 Matches: 30  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 2.10% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-893-238-3 (1-17056)

QY 337 LeuLysLeuProArgAlaSerHisLysAlaValAlaValAsnGlyAsnIleMetTrpValVal 356  
|||  
DB 2473 TTAAGCTTCCAGAGCCTCTCATAAAGCTGTGTCTCAATGAAATATATATGTTGGTGT 2414  
|||

QY 357 GlyGlyTyTrpMetPheAsnHisSerAspTy 366  
|||  
DB 2413 GCGGATATATGTTCAACCATTCAGATTAC 2384  
|||

RESULT 31  
US-10-197-824-1  
; Sequence 1, Application US/10197824  
; Publication No. US20040023219A1  
; GENERAL INFORMATION:  
; APPLICANT: ORIGENE TECHNOLOGIES INC  
; TITLE OF INVENTION: NOVEL PROSTATE CANCER GENES  
; FILE REFERENCE: 9U 106 R1  
; CURRENT APPLICATION NUMBER: US/10/197,824  
; CURRENT FILING DATE: 2002-07-19  
; NUMBER OF SEQ ID NOS: 39  
; SOFTWARE: PatentIn version 3.1

```

; SEQ ID NO 1
; LENGTH: 6733
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (575)..(4711)
; OTHER INFORMATION:
US-10-197-824-1

Alignment Scores:
Pred. No.: 8,31e-18 Length: 6733
Score: 29.00 Matches: 29
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.03% Indels: 0
DB: 16 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-197-824-1 (1-6733)

QY 1277 MetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuVal 1296
DB 4256 ATGGACCTTGTCAGTCTTTTTCACCTTCCTTCAGTGTTCCTATCCTTATTCGTGTG 4315

QY 1297 AlaAlaValValTropylsileysGln 1305
DB 4316 GCTGCTGTGTATGAGATCAACAA 4342

RESULT 32
US-10-027-632-264639
; Sequence 264639, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 264639
; LENGTH: 594
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-264639

Alignment Scores:
Pred. No.: 8,31e-18 Length: 594
Score: 28.00 Matches: 28
Percent Similarity: 100.00% Conservatives: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.96% Indels: 0
DB: 15 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-027-632-264639 (1-594)

QY 596 AlaCysAspArgTrpSerValLeuProArgProAspLeuHisAspValAsnArgPhe 615
DB 144 GCCTGTGACCGCTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATT 203

QY 616 GlyHisSerAlaValLeuHisAsn 623
DB 204 GGCCATTTCAGCAGTCTTTACACAAC 227

RESULT 34
US-10-027-632-270240
; Sequence 270240, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483

QY 596 AlaCysAspArgTrpSerValLeuProArgProAspLeuHisAspValAsnArgPhe 615
DB 144 GCCTGTGACCGCTGTCAGTGTCTCCAGACCTGATCTCCACCATGATGTCAACAGATT 203

QY 616 GlyHisSerAlaValLeuHisAsn 623
DB 204 GGCCATTTCAGCAGTCTTTACACAAC 227

RESULT 34
US-10-027-632-270240
; Sequence 270240, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
```

; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 270240  
; LENGTH: 629  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-270240

Alignment Scores:  
Pred. No.: 8,49e-16 Length: 629  
Score: 26.00 Matches: 26  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.82% Indels: 0  
DB: 15 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-027-632-270240 (1-629)

Qy 1189 GlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThr 1208  
Db 347 CAAACAGGGATTGGACATGTTTCATCAATGCTCCAGAAITTCACATCCACATCACC 406  
Qy 1209 TrpAlaIaSerPheSer 1214  
Db 407 TGGGCTGCCAGTTTCTCA 424

RESULT 35  
US-09-833-381-1916  
; Sequence 1916, Application US/09833381  
; Patent No. US20020132090A1  
; GENERAL INFORMATION:  
; APPLICANT: Robison, Keith E.  
; TITLE OF INVENTION: No. US20020132090A1el Nucleic Acid and Protein Homologs  
; FILE REFERENCE: 5800-119  
; CURRENT APPLICATION NUMBER: US/09/833,381  
; CURRENT FILING DATE: 2001-04-11  
; PRIOR APPLICATION NUMBER: 09/516,448  
; PRIOR FILING DATE: 2000-02-29  
; NUMBER OF SEQ ID NOS: 2050  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1916  
; LENGTH: 415  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1)...(415)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-833-381-1916

Alignment Scores:  
Pred. No.: 5.38e-12 Length: 415  
Score: 22.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.54% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-833-381-1916 (1-415)

Qy 222 ThrSerGlyTrpAlaLeuLeuHisPhePheSerAspAlaAlaIaTyAsnLeuThrGlyPhe 241  
Db 247 ACATCGGCTATGCATGTTACATTTTTTATGATGCTGCGTATTAATCACTAATCGTTTC 306

Qy 242 AsnIle 243  
Db 307 AACATT 312

## RESULT 36

US-10-311-455-170/c  
; Sequence 170, Application US/10311455  
; Publication No. US20030143606A1  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation of Cytosine  
; FILE REFERENCE: 5013.1014  
; CURRENT APPLICATION NUMBER: US/10/311,455  
; CURRENT FILING DATE: 2002-12-16  
; PRIOR APPLICATION NUMBER: PCT/EP01/07537  
; PRIOR FILING DATE: 2001-07-02  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 2424  
; SEQ ID NO 170  
; LENGTH: 5312  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
US-10-311-455-170

## Alignment Scores:

Pred. No.: 5.93e-11 Length: 5312  
Score: 22.00 Matches: 22  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 1.54% Indels: 0  
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-311-455-170 (1-5312)

Qy 47 LeuArgLeuProArgLeuLeuSerProLeuArgProArgLeuLeuLeuLeu 66  
Db 174 CTACGCTCCGCGACTACTATCTCCACCGTACGACCACTACTACTACTATTA 115  
Qy 67 LeuLeu 68  
Db 114 TTACTC 109

## RESULT 37

US-09-908-975-15421  
; Sequence 15421, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: WASSERMAN, Alon  
; APPLICANT: MINTZ, Eli  
; APPLICANT: MINTZ, Liat  
; APPLICANT: FAIGLER, Simchon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPlicing  
; FILE REFERENCE: 36688-0005  
; CURRENT APPLICATION NUMBER: US/09/908,975  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/287,724  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: US 60/221,607  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 32337  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 15421  
; LENGTH: 60



```
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-15421

Alignment Scores:
Pred. No.:      8,43e-11      Length:      60
Score:          20.00         Matches:     20
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:    1.40%        Indels:         0
DB:             10          Gaps:           0

US-09-787-097-12 (1-1429) x US-09-908-975-15421 (1-60)
QY 1183 ValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsn 1202
Db 1 GTGGCTACTCTCTGACGACAAACAAACAGGGATTTGGACATGTTTCATCAATGCCTCCAAGAAT 60

RESULT 38
US-09-908-975-29990
; Sequence 29990, Application US/09908975
; Publication No. US20030165843A1
; GENERAL INFORMATION:
; APPLICANT: SHOSHAN, Avi
; APPLICANT: WASSERMAN, Alon
; APPLICANT: MINTZ, Eli
; APPLICANT: MINTZ, Liat
; APPLICANT: FAIGLER, Simchon
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICE
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; PRIOR FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 29990
; LENGTH: 65
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-908-975-29990

Alignment Scores:
Pred. No.:      8.8e-09      Length:      65
Score:          18.00         Matches:     18
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:    1.26%        Indels:         0
DB:             10          Gaps:           0

US-09-787-097-12 (1-1429) x US-09-908-975-29990 (1-65)
QY 946 SerSerGluCysMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAla 963
Db 10 AGCTCGAGTGCATGTGGTGCAGTAACATCAACAGCAGTGTGGACTTCCATGCC 63

RESULT 39
US-09-728-446-450
; Sequence 450, Application US/09728446
; Patent No. US20020081568A1
; GENERAL INFORMATION:
; APPLICANT: Friedreich, Glenn
; APPLICANT: Zambrowicz, Brian
; APPLICANT: Sands, Arthur T.
; TITLE OF INVENTION: No. US20020081568A1el Murine Polynucleotide Sequences
; FILE REFERENCE: LEX-0101-USA
; CURRENT APPLICATION NUMBER: US/C9/728,446
; CURRENT FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/168,270

; PRIOR FILING DATE: 1999-12-01
; NUMBER OF SEQ ID NOS: 1461
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 450
; LENGTH: 233
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE: misc feature
; NAME/KEY: (1)...(233)
; LOCATION: (1)...(233)
; OTHER INFORMATION: n = A,T,C or G
US-09-728-446-450

Alignment Scores:
Pred. No.:      2.83e-06      Length:      233
Score:          16.00         Matches:     16
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:    1.12%        Indels:         0
DB:             9          Gaps:           0

US-09-787-097-12 (1-1429) x US-09-728-446-450 (1-233)
QY 62 LeuLeuLeuLeuLeuLeuProProLeuLeuLeuLeuLeuLeu 77
Db 60 CTCCTGCTGCTGCTGCTGCTGCCCTCCGCCCTGCTACTGCTGCTGTTG 107

RESULT 40
US-09-925-300-486/c
; Sequence 486, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruben,
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 486
; LENGTH: 2572
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE: misc feature
; NAME/KEY: (823)
; LOCATION: (823)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-300-486

Alignment Scores:
Pred. No.:      0.00263      Length:      2572
Score:          14.00         Matches:     14
Percent Similarity: 100.00%   Conservative: 0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:    0.98%        Indels:         0
DB:             9          Gaps:           0

US-09-787-097-12 (1-1429) x US-09-925-300-486 (1-2572)
QY 64 LeuLeuLeuLeuLeuProProLeuLeuLeuLeuLeuLeu 77
Db 60 CTGCTGCTGCTGCTGCCACCGCGCTGCTGCTGCTGCTGCTG 19

RESULT 41
US-10-198-846-13651/c
; Sequence 13651, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
```



```
US-10-465-572-17/c
; Sequence 17, Application US/10465572
; Publication No. US20030207840A1
; GENERAL INFORMATION:
; APPLICANT: Riggins, Gregory
; APPLICANT: Lal, Anita
; TITLE OF INVENTION: GENES INDUCED BY HYPOXIA
; FILE REFERENCE: 000250.00012
; CURRENT APPLICATION NUMBER: US/10/465,572
; CURRENT FILING DATE: 2003-06-20
; PRIOR APPLICATION NUMBER: US/10/201,642
; PRIOR FILING DATE: 2002-07-24
; PRIOR APPLICATION NUMBER: 60/307,600
; PRIOR FILING DATE: 2001-07-26
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-465-572-17

Alignment Scores:
Pred. No.: 0.00389      Length: 3901
Score: 14.00           Matches: 14
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.98%              Indels: 0
DB: 15                        Gaps: 0

US-09-787-097-12 (1-1429) x US-10-465-572-17 (1-3901)
QY 64 LeuLeuLeuLeuProProLeuLeuLeuLeuLeu 77
Db 80 CTGCTGCTGCTGTCACCGCGCTGCTGCTGCTG 39

RESULT 45
US-10-159-563-292/c
; Sequence 292, Application US/10159563
; Publication No. US20040009154A1
; GENERAL INFORMATION:
; APPLICANT: Khan, Javed
; APPLICANT: Ringner, Markus
; APPLICANT: Peterson, Carsten
; APPLICANT: Meltzer, Paul
; TITLE OF INVENTION: SELECTIONS OF GENES AND METHODS OF USING THE SAME FOR
; TITLE OF INVENTION: DIAGNOSIS AND FOR TARGETING THE THERAPY OF SELECT CANCERS
; FILE REFERENCE: 11613.56US11
; CURRENT APPLICATION NUMBER: US/10/159,563
; CURRENT FILING DATE: 2002-12-09
; PRIOR APPLICATION NUMBER: US 10/133,937
; PRIOR FILING DATE: 2002-04-25
; NUMBER OF SEQ ID NOS: 444
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 292
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-159-563-292

Alignment Scores:
Pred. No.: 0.00389      Length: 3901
Score: 14.00           Matches: 14
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.98%              Indels: 0
DB: 15                        Gaps: 0

US-09-787-097-12 (1-1429) x US-10-159-563-292 (1-3901)
QY 64 LeuLeuLeuLeuProProLeuLeuLeuLeuLeu 77
Db 80 CTGCTGCTGCTGTCACCGCGCTGCTGCTGCTG 39
```

```
RESULT 46
US-10-372-683-40/c
; Sequence 40, Application US/10372683
; Publication No. US20040009171A1
; GENERAL INFORMATION:
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Peale Jr., Franklin V.
; APPLICANT: Wu, Thomas D.
; TITLE OF INVENTION: METHODS FOR THE TREATMENT OF CARCINOMA
; FILE REFERENCE: P1928R1P1
; CURRENT APPLICATION NUMBER: US/10/372,683
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 10/271,690
; PRIOR FILING DATE: 2002-10-16
; PRIOR APPLICATION NUMBER: US 60/344,534
; PRIOR FILING DATE: 2001-10-18
; NUMBER OF SEQ ID NOS: 49
; SEQ ID NO 40
; LENGTH: 3901
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-372-683-40

Alignment Scores:
Pred. No.: 0.00389      Length: 3901
Score: 14.00           Matches: 14
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%    Mismatches: 0
Query Match: 0.98%              Indels: 0
DB: 15                        Gaps: 0

US-09-787-097-12 (1-1429) x US-10-372-683-40 (1-3901)
QY 64 LeuLeuLeuLeuProProLeuLeuLeuLeuLeu 77
Db 80 CTGCTGCTGCTGTCACCGCGCTGCTGCTGCTG 39

RESULT 47
US-10-175-523-57/c
; Sequence 57, Application US/10175523
; Publication No. US20030096264A1
; GENERAL INFORMATION:
; APPLICANT: Brockman, Jeffrey
; APPLICANT: Evans, David
; APPLICANT: Hook, Derek
; APPLICANT: Klimczak, Leszek
; APPLICANT: Laeng, Pascal
; APPLICANT: Palfreyman, Michael
; APPLICANT: Rajan, Prithi
; TITLE OF INVENTION: MULTI-PARAMETER HIGH THROUGHPUT SCREENING ASSAYS (MPHTS)
; FILE REFERENCE: 3235/1J795-US3
; CURRENT APPLICATION NUMBER: US/10/175,523
; CURRENT FILING DATE: 2002-06-18
; PRIOR APPLICATION NUMBER: US 60/299,151
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: US 60/317,828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: US 60/325,150
; PRIOR FILING DATE: 2001-09-25
; PRIOR APPLICATION NUMBER: US 60/333,047
; PRIOR FILING DATE: 2001-11-14
; PRIOR APPLICATION NUMBER: US 60/349,936
; PRIOR FILING DATE: 2002-01-18
; PRIOR APPLICATION NUMBER: US 60/361,834
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 197
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 57
; LENGTH: 225863
; TYPE: DNA
; ORGANISM: Mus musculus
US-10-175-523-57
```



US-10-108-260A-876  
Alignment Scores:

```

; APPLICANT: PATTERSON, Chandra
; TITLE OF INVENTION: MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-0731-1 USA

```

; CURRENT APPLICATION NUMBER: US/09/969,680A  
; CURRENT FILING DATE: 2001-10-02  
; PRIOR APPLICATION NUMBER: US00/22315  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/149,641  
; PRIOR FILING DATE: 1999-08-17  
; PRIOR APPLICATION NUMBER: 60/164,203  
; PRIOR FILING DATE: 1999-11-09  
; NUMBER OF SEQ ID NOS: 74  
; SOFTWARE: PERL Program  
; SEQ ID NO 39  
; LENGTH: 2656  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20030124649A1 997947CB1  
; FEATURE:  
; LOCATION: (1) ... (2656)  
; OTHER INFORMATION: a, t, c, g, or other  
US-09-969-680A-39

Alignment Scores:  
Pred. No.: 0.0267 Length: 2656  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 10 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-969-680A-39 (1-2656)  
Qy 61 LeuLeuLeuLeuLeuLeuLeuProProLeuLeu 73  
Db 86 CTGCTGTATTGCTACTGCTGCTGCGCGCGCTCTGCTT 124

RESULT 55  
US-10-301-822-3  
; Sequence 3, Application US/10301822  
; Publication No. US20030148410A1  
; GENERAL INFORMATION:  
; APPLICANT: Millennium Pharmaceuticals, Inc.  
; APPLICANT: Berger, Allison  
; APPLICANT: Guillemette, Tracy L.  
; APPLICANT: Kamatkar, Shubhangi  
; APPLICANT: Schlegel, Robert  
; APPLICANT: Monahan, John B.  
; APPLICANT: Thibodeau, Stephen N.  
; APPLICANT: Burgart, Lawrence J.  
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND  
; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
; TITLE OF INVENTION: THERAPY OF COLON CANCER  
; FILE REFERENCE: MP01-029P2RM  
; CURRENT APPLICATION NUMBER: US/10/301,822  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 60/339,971  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: US 60/361,978  
; PRIOR FILING DATE: 2002-03-05  
; PRIOR APPLICATION NUMBER: US 60/381,988  
; PRIOR FILING DATE: 2002-05-20  
; NUMBER OF SEQ ID NOS: 228  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 6691  
; TYPE: DNA  
; ORGANISM: Homo Sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (82) ... (3717)  
US-10-301-822-3

Alignment Scores:  
Pred. No.: 0.0637 Length: 6691  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 14 Gaps: 0

US-09-787-097-12 (1-1429) x US-10-301-822-3 (1-6691)  
Qy 61 LeuLeuLeuLeuLeuLeuLeuProProLeuLeu 73  
Db 127 CTGCTGCTGCTGCTGCTGCTGCGCGCGCTCTG 165

RESULT 56  
US-09-788-711A-3  
; Sequence 3, Application US/09788711A  
; Patent No. US20020058328A1  
; GENERAL INFORMATION:  
; APPLICANT: Tania Tamsin Testa  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GP-30225  
; CURRENT APPLICATION NUMBER: US/09/788,711A  
; CURRENT FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 0004196.2  
; PRIOR FILING DATE: 2000-02-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 3  
; LENGTH: 8772  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
US-09-788-711A-3

Alignment Scores:  
Pred. No.: 0.0821 Length: 8772  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.91% Indels: 0  
DB: 9 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-788-711A-3 (1-8772)  
Qy 61 LeuLeuLeuLeuLeuLeuLeuProProLeuLeu 73  
Db 52 CTGCTGCTGTGCTGCTGCTGCGCGCGCTATTG 90

RESULT 57  
US-09-788-711A-1  
; Sequence 1, Application US/09788711A  
; Patent No. US20020058328A1  
; GENERAL INFORMATION:  
; APPLICANT: Tania Tamsin Testa  
; TITLE OF INVENTION: NOVEL COMPOUNDS  
; FILE REFERENCE: GP-30225  
; CURRENT APPLICATION NUMBER: US/09/788,711A  
; CURRENT FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 0004196.2  
; PRIOR FILING DATE: 2000-02-19  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 8871  
; TYPE: DNA  
; ORGANISM: HOMO SAPIENS  
US-09-788-711A-1

Alignment Scores:  
Pred. No.: 0.083 Length: 8871  
Score: 13.00 Matches: 13  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0

```

; TITLE OF INVENTION: ANTIGENIC PEPTIDES AND ANTIBODIES FOR G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 1920-4-4
; CURRENT APPLICATION NUMBER: US/10/225,567A
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/257,144
; PRIOR FILING DATE: 2000-12-19
; NUMBER OF SEQ ID NOS: 2292
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 523
; LENGTH: 10531
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-225-567A-523

Alignment Scores:
Pred. No.:      0.0976      Length:      10531
Score:          13.00      Matches:       13
Percent Similarity: 100.00%  Conservative:  0
Best Local Similarity: 100.00%  Mismatches:   0
Query Match:    0.91%      Indels:       0
DB:             14        Gaps:         0

US-09-787-097-12 (1-1429) x US-10-225-567A-523 (1-10531)

QY      61 LeuLeuLeuLeuLeuLeuLeuLeuProProLeuLeu 73
      |||||
Db      114 CTGCTGCTTTGCTGCTGCTGCTGCCGCCCAATTG 152

RESULT 60
US-10-292-798-931
; Sequence 931, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ASURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084935/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18
; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 931
; LENGTH: 24370
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; LOCATION: source
; FEATURE:
; LOCATION: (1)..(24370)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (201)..(3510)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (8553)..(9200)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11163)..(11385)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11634)..(11749)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11929)..(12017)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (12408)..(12565)

```



















GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: March 1, 2004, 07:13:00 ; Search time 236 Seconds

(without alignments)  
3360.275 Million cell updates/sec

Title: US-09-787-097-12

Perfect score: 1429

Sequence: 1 MVAAAAATEARLRRRTAATA.....KSGAVNRKQPPAOPGTCTI 1429

Scoring table:

OLIGO  
Xgapop 60.0 , Xgapext 60.0  
Ygapop 60.0 , Ygapext 60.0  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues

Word size: 1

Total number of hits satisfying chosen parameters: 1360453

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 100 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp  
-Q/cgn2\_1/USPTO\_spool\_p/US09787097/runat\_26022004\_063015\_23781/app\_query.fasta\_1.1607  
-DB=issued Patents NA -QFMT=fastap -SUFFIX=oligo.rni -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=Bits -START=1 -END=1 -MATRIX=oligo -TRANS=human4.cdi  
-LIST=100 -DOCALLIGN=200 -THR SCORE=quality -THR MIN=1 -ALIGN=75 -MODE=LOCAL  
-OUTFMT=ptc -NORV=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USPR=US09787097 @cgn 1.118 @runat\_26022004\_063015\_23781 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN\_TIMEOUT=30 -THREADS=1 -XGAPOP=60 -XGAPEXT=60 -FGAPOP=6  
-FGAPEXT=7 -YGAPEXT=60 -YGAPEXT=60 -DELOP=6 -DELEXT=7

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/2/ina/5A.COMB.seq.\*  
2: /cgn2\_6/ptodata/2/ina/5B.COMB.seq.\*  
3: /cgn2\_6/ptodata/2/ina/6A.COMB.seq.\*  
4: /cgn2\_6/ptodata/2/ina/6B.COMB.seq.\*  
5: /cgn2\_6/ptodata/2/ina/PCTUS.COMB.seq.\*  
6: /cgn2\_6/ptodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1141	79.8	8589	3	US-09-245-041-14
2	979	68.5	4072	3	US-09-245-041-16
3	503	35.2	6370	3	US-09-245-041-12
4	496	34.7	2625	3	US-09-245-041-18
5	206	14.4	8827	3	US-09-245-041-1
6	94	6.6	2419	3	US-09-245-041-8
C 7	82	5.7	246	4	US-09-702-705-966
C 8	82	5.7	246	4	US-09-736-457-966
C 9	82	5.7	246	4	US-09-614-1248-966
C 10	82	5.7	246	4	US-09-671-325-966
11	76	5.3	1051	3	US-09-245-041-10
12	73	5.1	90050	3	US-09-245-041-5
					Sequence 14, Appl
					Sequence 16, Appl
					Sequence 18, Appl
					Sequence 1, Appl
					Sequence 8, Appl
					Sequence 966, App
					Sequence 966, App
					Sequence 966, App
					Sequence 10, Appl
					Sequence 5, Appl

C 13	51	3.6	226	3	US-09-245-041-6	Sequence 6, Appl
C 14	45	3.1	221	3	US-09-245-041-7	Sequence 7, Appl
C 15	33	2.3	5973	3	US-09-245-041-3	Sequence 4, Appl
C 16	30	2.1	17056	3	US-09-245-041-3	Sequence 3, Appl
C 17	22	1.5	415	4	US-09-833-381-1916	Sequence 1916, Ap
C 18	13	0.9	429	3	US-08-905-223-61	Sequence 61, Appl
C 19	13	0.9	2023	4	US-09-431-522-6	Sequence 6, Appl
C 20	13	0.9	2450	4	US-09-431-522-2	Sequence 2, Appl
C 21	13	0.9	6692	4	US-09-431-522-1	Sequence 1, Appl
C 22	12	0.8	1103	3	US-08-927-219-54	Sequence 54, Appl
C 23	12	0.8	2426	3	US-09-028-327-2	Sequence 2, Appl
C 24	12	0.8	2426	4	US-09-571-078A-2	Sequence 2, Appl
C 25	11	0.8	639	4	US-09-336-536-57	Sequence 57, Appl
C 26	11	0.8	1432	4	US-09-620-312D-420	Sequence 420, App
C 27	11	0.8	1739	4	US-09-148-545-11	Sequence 11, Appl
C 28	11	0.8	1858	4	US-09-336-536-56	Sequence 56, Appl
C 29	11	0.8	1899	4	US-09-919-060-15	Sequence 15, Appl
C 30	11	0.8	1899	4	US-09-919-060-16	Sequence 16, Appl
C 31	11	0.8	2362	4	US-09-907-794A-194	Sequence 194, App
C 32	11	0.8	2362	4	US-09-905-125A-194	Sequence 194, App
C 33	11	0.8	2362	4	US-09-902-775A-194	Sequence 194, App
C 34	11	0.8	2520	1	US-08-405-254-9	Sequence 9, Appl
C 35	11	0.8	2693	4	US-09-919-060-12	Sequence 12, Appl
C 36	11	0.8	2693	4	US-09-919-060-14	Sequence 14, Appl
C 37	11	0.8	2784	1	US-08-261-465-3	Sequence 3, Appl
C 38	11	0.8	2784	1	US-08-405-254-7	Sequence 7, Appl
C 39	11	0.8	3110	4	US-09-023-655-1072	Sequence 1072, Ap
C 40	11	0.8	3132	2	US-08-224-482-3	Sequence 3, Appl
C 41	11	0.8	3132	3	US-09-205-921-1	Sequence 1, Appl
C 42	11	0.8	3132	4	US-09-300-958A-32	Sequence 32, Appl
C 43	11	0.8	3132	4	US-09-702-705-320	Sequence 320, App
C 44	11	0.8	3132	4	US-09-736-457-320	Sequence 320, App
C 45	11	0.8	3132	4	US-09-614-124B-320	Sequence 320, App
C 46	11	0.8	3132	4	US-09-671-325-320	Sequence 320, App
C 47	11	0.8	3132	4	US-09-589-184-320	Sequence 320, App
C 48	11	0.8	3200	4	US-08-884-569A-1	Sequence 1, Appl
C 49	11	0.8	3311	4	US-09-367-891A-5	Sequence 5, Appl
C 50	11	0.8	4169	4	US-09-166-350-32	Sequence 32, Appl
C 51	11	0.8	5828	4	US-09-023-655-1386	Sequence 1386, Ap
C 52	11	0.8	392000	4	US-10-027-983-11	Sequence 11, Appl
C 53	10	0.7	109	2	US-08-811-492-132	Sequence 132, App
C 54	10	0.7	109	2	US-08-811-492-133	Sequence 133, App
C 55	10	0.7	122	1	US-08-478-312-55	Sequence 55, Appl
C 56	10	0.7	122	1	US-08-485-302-55	Sequence 55, Appl
C 57	10	0.7	122	1	US-08-476-169-60	Sequence 60, Appl
C 58	10	0.7	122	1	US-08-484-083-60	Sequence 60, Appl
C 59	10	0.7	122	2	US-08-764-640-237	Sequence 237, App
C 60	10	0.7	122	3	US-09-973-225-225	Sequence 225, App
C 61	10	0.7	122	3	US-09-244-298A-237	Sequence 237, App
C 62	10	0.7	122	3	US-09-516-704-237	Sequence 237, App
C 63	10	0.7	122	4	US-09-549-030-225	Sequence 225, App
C 64	10	0.7	122	4	US-09-832-230A-237	Sequence 237, App
C 65	10	0.7	577	3	US-08-938-548B-21	Sequence 21, Appl
C 66	10	0.7	577	4	US-08-939-093A-21	Sequence 21, Appl
C 67	10	0.7	577	4	US-09-211-932C-21	Sequence 21, Appl
C 68	10	0.7	614	4	US-09-325-932A-31	Sequence 31, Appl
C 69	10	0.7	867	4	US-09-833-381-782	Sequence 782, App
C 70	10	0.7	999	4	US-08-811-481-26	Sequence 26, Appl
C 71	10	0.7	999	4	US-09-876-527-26	Sequence 26, Appl
C 72	10	0.7	1011	4	US-09-876-527-27	Sequence 27, Appl
C 73	10	0.7	1011	4	US-09-876-527-27	Sequence 27, Appl
C 74	10	0.7	1146	4	US-09-489-039A-3636	Sequence 3636, Ap
C 75	10	0.7	1276	3	US-09-177-325-2	Sequence 2, Appl
C 76	10	0.7	1276	3	US-09-411-812A-2	Sequence 2, Appl
C 77	10	0.7	1425	1	US-09-590-113-2	Sequence 1, Appl
C 78	10	0.7	1425	1	US-08-464-148-1	Sequence 1, Appl
C 79	10	0.7	1425	1	US-08-385-500-1	Sequence 1, Appl
C 80	10	0.7	1425	1	US-08-846-784-1	Sequence 1, Appl
C 81	10	0.7	1436	4	US-09-471-395-2	Sequence 2, Appl
C 82	10	0.7	1436	4	US-09-788-345-11	Sequence 11, Appl
C 83	10	0.7	1537	2	US-08-833-008-1	Sequence 1, Appl
C 84	10	0.7	1578	4	US-09-130-491-10	Sequence 10, Appl
C 85	10	0.7	1768	4	US-09-833-381-523	Sequence 523, App

86 10 0.7 1895 4 US-09-023-655-1162 Sequence 1162, Ap  
87 10 0.7 1970 3 US-08-938-548B-1 Sequence 1, Appli  
88 10 0.7 1970 4 US-08-939-093A-1 Sequence 1, Appli  
89 10 0.7 1970 4 US-09-211-823C-1 Sequence 1, Appli  
90 10 0.7 2073 3 US-08-816-977-46 Sequence 46, Appl  
91 10 0.7 2073 4 US-09-334-477-46 Sequence 46, Appl  
92 10 0.7 2085 3 US-08-816-977-48 Sequence 48, Appl  
93 10 0.7 2085 3 US-09-334-477-48 Sequence 48, Appl  
94 10 0.7 2127 3 US-08-816-977-32 Sequence 32, Appl  
95 10 0.7 2127 4 US-09-334-477-32 Sequence 32, Appl  
96 10 0.7 2136 3 US-08-816-977-34 Sequence 34, Appl  
97 10 0.7 2136 4 US-09-334-477-34 Sequence 34, Appl  
98 10 0.7 2329 1 US-08-455-559-9 Sequence 9, Appli  
99 10 0.7 2329 3 US-09-145-060-9 Sequence 9, Appli  
100 10 0.7 2329 5 PCT-US94-00657-9 Sequence 9, Appli

## ALIGNMENTS

RESULT 1  
US-09-245-041-14  
; Sequence 14, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, K.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 14  
; LENGTH: 8589  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-14

Alignment Scores:  
Pred. No.: 0 Length: 8589  
Score: 1141.00 Matches: 1341  
Percent Similarity: 99.70% Conservative: 0  
Best Local Similarity: 99.70% Mismatches: 2  
Query Match: 79.85% Indels: 4  
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-245-041-14 (1-8589)

QY 87 AlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
DB 256 GCGCGCGGTGTCGGTCCGCGCAGCAGCGAGGCGGAGGATGACCGCGGTGTC 315  
QY 107 AsnGlyArgCysAsnProGlyThrGlyGlnCysValCysProAlaGlyTrpValGly 126  
DB 316 AACGGCGGTGCTGCAACCTCGCAGCGGCGGCGGCGGCGGCGGCGGCGGCGG 375  
QY 127 GluGlnCysGlnHisCysGlyGlyArgPheArgLeuThrGlySerGlyPheValThr 146  
DB 376 GAGCAATGCCAGCACTCGCGGGGCGGCTTCAGACTAACTGATCTTCTGGGTTTGACA 435  
QY 147 AspGlyProGlyAsnTrpLysTrpLysThrLysCysThrTrpLeuLeuGluGlnPro 166  
DB 436 GATGGACCTGCAATTAATAACAAACGAGTGCAGCTGCTGCTGCTGCTGCTGCTG 495  
QY 167 AsnArgLeuMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186  
DB 496 AATAGAAATATGAGATTCGTTTCATCATTTTCTACAGAGTGTAGTTGGGACCATTTA 555

QY 187 TyrValTyrAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206  
DB 556 TATGTTTATGATGGGAGCTCAATTTATGCACCGCTAGTTGCTGCTATTTAGTGGCTCAT 615  
QY 207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226  
DB 616 GTTCTCTGAGAGAGATGGCAATGAGACTGTCCTCAGGTGTTGTCACATCAGGTATGACC 675  
QY 227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTyrSer 246  
DB 676 TTGCTGCAATTTTATGATGCTGCTTATATTTGACTGATTAATTAATTAATTAATTAAT 735  
QY 247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266  
DB 736 TTTGATATGTCTCCAAATACTGCTCAGGCCGAGAGAGTGTAAAGATCAGTAATAGCAGC 795  
QY 267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286  
DB 796 GA-TACTGTTGAATGTGAATGTTCTGAAACTGGAAAGGTTGAAGCATGTGACATTCCTCA 854  
QY 286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValaArgG 306  
DB 855 CTGTACAGACCACTGTGGTTCCTCATCGAGGCACTGCAATTCAGTGATGTGACAGG 914  
QY 306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnG 326  
DB 915 ATGCTCTCTCTCAGACTGCGAGGCTCTGAGTGTTCAGTTCCTCTGACGAGTAACCA 974  
QY 326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346  
DB 975 GTCATTTTGGACTCGAGAGGAATATTTCACTTAAGCTCCCGAGAGCATCTCTAAGC 1034  
QY 346 aValValAsnGlyAsnIleMetTrpValValGlyClyTyrMetPheAsnHisSerAspTy 366  
DB 1035 TGTGTCATGTAAGCAATTAATGTCGGTGTGTTGGAGGATATATGTTCAACCACTCAGAT 1094  
QY 366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386  
DB 1095 TAAATGCTTCTAGCGTATGACCTGCTCTAGGAGTGGCTTCCATTAACCGCTCTGT 1154  
QY 386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406  
DB 1155 GAACAAATGCTGTTGTAGATATGTCATCTTTGGCATTTATACAAGGATAAAATTTACAT 1214  
QY 406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426  
DB 1215 GTATGGAGGAAAAATTTGATTC-AACTGGGAATGTGACCAATGAGTTGAGAGTTTTTCA 1273  
QY 426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446  
DB 1274 TTCATATGATCATGGGTGTTGTGACCCCTAAGGCAAGGAGCATGTATGACGTGGTTG 1333  
QY 446 lyHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466  
DB 1334 GGCACCTGCAACATTTTACACTGAAGAATGCGGAGTGGTCACTGCTGCTCATCTTTG 1393  
QY 466 lyHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsn 486  
DB 1394 GTCACTGCGCTCTCTATGGATATATAGCAATGTGAGGAATATGATTTGGATTAAGAACA 1453  
QY 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSery 506  
DB 1454 CATGGAGTATATTACACACCAGGCTGCTTGTGCAAGGGGTACGGGCATGACGAGT 1513  
QY 506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyTyrLysAlaPheSerAlaA 526  
DB 1514 TTTACGACCATAGACAGGCGCTTACGTTCAAGTGGCTACAGGCTTTTCACTGCCA 1573  
QY 526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546  
DB 1574 ATAAGTACCGCTTGCAGATGATCTCTACCGATATGATGTGGATACCCAGATGTGGACCA 1633  
QY 546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566

Db |||||TTCTTAAGGACAGCCGATTTTCCGTTACTTGTGCACACAGCTGTGTAGTGAAGTGAACCA 1693  
Qy 566 etLeuValPheGlyGlyAenThrHisAsnAspThrSerMetSerHisGlyAlaLeuCyseP 586  
Db 1694 TGCTGGTGTGTGGGGAAACACACACAATGACACATCTATAGCCATGGCCCAATGCT 1753  
Qy 586 heSerSerAspPheMetAlaTyraSpilleAlaCysAspArgTrpSerValLeuProArgP 606  
Db 1754 TCCTCTCAGATTTTCAGCCCTATGACATTCGCTGTGACCGCTGGTCAGTCTTCCCGAC 1813  
Qy 606 roAspLeuHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrm 626  
Db 1814 CTGATCTCCACCACATGATCAACAGATTTGGCCATTCAGCAGTCTTACACACAGACCA 1873  
Qy 626 etTyraValPheGlyGlyPheAsnSerLeuLeuLeuSerAspIleLeuValPheThrSerG 646  
Db 1874 TGTATGTGTTCGGTGTTCATATAGTCTCTCTCAGCAGCATCTCTGGTATTACCTCGG 1933  
Qy 646 luGlnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
Db 1934 AACAGTGTGATGGCCATCGAGTGAAGCCGCTTTTGTAGCAGCAGGACCTGTGTATTCGGT 1993  
Qy 666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686  
Db 1994 GTGTGTGGAACACAGAGTCTCTCAGTGTATCTCTGTGGGGCTGGCAACTGATGAACAG 2053  
Qy 686 luGluLysLeuLysSerGluCysePheSerLysArgThrLeuAspHisAspArgCyseAspG 706  
Db 2054 AAGAAAAGTTAAAAATCAGAAATGTTTTTCCAAAAGAACTCTTGACCATGACAGATGTGACC 2113  
Qy 706 lnhIstrAspCysTyraSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726  
Db 2114 AGCACACAGATTTTACAGCTGCACAGCCCAACCAATGACTGCCACTGGTGAATGACC 2173  
Qy 726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyraGluA 746  
Db 2174 ATTGTGTCCCGGAACCAACAGCTGCTCAGAGGCCAGATCTCCATTTTAGTATGAGA 2233  
Qy 746 snCysProLysAspAsnProMetTyraTyraCysAsnLysLysThrSerCysArgSerCysA 766  
Db 2234 ATTGCCCCAGAGTAAACCTTATCTACTGTAAACAGAGACAGCTGCAGAGAGCTGTG 2293  
Qy 766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnCysIleAlaLeuProGluA 786  
Db 2294 CCCTGGACCAAGAACTGCCAGTGGAGCCCGGAATCAGAGTGCATTTGCCCTGCCGAAA 2353  
Qy 786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuLysIleThrThrAlaL 806  
Db 2354 ATATCTGTGCATTGGCTGGCAATTGGTTGGAAACTCATGTTGAAAATTTACTTACTGCCA 2413  
Qy 806 ysGluAsnTyraAspAsnAlaLysLeuPheCysArgAsnHisAsnAlaLeuAlaSerL 826  
Db 2414 AGGAGAATTATGACAATGTCTAAATTTGTCTGTAGGAACCAATGCCCTTTTGGCTTCTC 2473  
Qy 826 euThrThrGlnLysLysValGluPheValLeuLysGlnLeuArgIleMetGlnSerSerG 846  
Db 2474 TTACAACCCAGAGAAGTGAATTTGTCTTAAGCAGCTGCGCAATAATGACAGTCATCTC 2533  
Qy 846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgLysIleAsnValSerT 866  
Db 2534 AGAGCATGTCCAAGCTCACCTTAACCCATGGGTGGCCCTTCGGAAGATCAATGTGTCTCT 2593  
Qy 866 yrTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886  
Db 2594 ACTGGTCTGGGAAGATATGTCCCATTTTACAAATAGTTTACTACAGTGGATGCCCTCTG 2653  
Qy 886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysA 906  
Db 2654 AGCCAGTGTGCTGGGATTTCTGTGGAATTTTATCAGAACCCAGTACTCTCGGGACTGAAG 2713  
Qy 906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926

Db 2714 CTGCAACCTGCATCAACCCACTCAATGGTAGTGTCTGTGAAAGGCGCTGCAAAACACACAGTG 2773  
Qy 926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlys 946  
Db 2774 CTNAGCAGTGGCGGACACCATGTGCTTGAGGACAGCATGTGGAGATGTGCACACGCGCA 2833  
Qy 946 erSerGluCyseMetTrpCysSerAsnMetLysGlnCysValAspSerAsnAlaTyraValA 966  
Db 2834 GCTCTAGTGCATGTGTGCAGCAACATGAAGCAGTGTGTGGACTTCCAATGCCATGTGG 2893  
Qy 966 laSerPheProPheGlyGlnCysMetGluTrpTyraThrMetSerThrCysProProGluA 986  
Db 2894 CCTCTTCCCTTTGGCCAGTGTATGGAATGTTATACGATGAGCACCTGCCCCCTGAAA 2953  
Qy 986 snCysSerGlyTyraCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006  
Db 2954 ATTGTTCAGGCTACTGTACTCTGTAGTCTATTGCTTGGAGCAACAGGCTGTGGCTGTGTA 3013  
Qy 1006 hrAspProSerAsnThrGlyLysGlyCysIleGluGlySerTyraLysGlyProValL 1026  
Db 3014 CTGATCCCGACCAATACTGGCAAGGGAATGATAGAGGGTTCCTATTAAGAGCACAGTGA 3073  
Qy 1026 ysMetProSerGlnAlaProThrGlyAsnPheTyraProGlnProLeuLeuAsnSerSerM 1046  
Db 3074 AGATGCCCTTCGAAAGCCCTTACAGGAATTTCTATCCACAGCCCTGCTCAATCCAGCA 3133  
Qy 1046 etCysLeuGluAspSerArgTyraAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066  
Db 3134 TGTGTCTAGAGACAGCAGATACAACCTGGTCTTTCTTCTTCTCTGCTCCAGCTTGGCAATGCA 3193  
Qy 1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrThrg 1086  
Db 3194 ACGGCCACAGTAAATGATCATCAATCAGACATCTGTGAGAACTGTGAGAACCTGACCACAG 3253  
Qy 1086 lyLysHisCysGluThrCysIleSerGlyPheTyraGlyAspProThrAsnGlyGlyLysC 1106  
Db 3254 GCAAGCACTGGAGACCTGCATATCTGGCTTCTACGGTGATCCCAATGGAGGGAAT 3313  
Qy 1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126  
Db 3314 GTACCCCATGCAAGTGCAATGGGACCGCTCTGTGTGCAACCAACACGCGGCAAGTCT 3373  
Qy 1126 heCysThrLysGlyValLysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146  
Db 3374 TCTGCACCAACCAAGGCGTCAAGGGGACGAGTSCAGCTATGTGAGTGAATAATCGAT 3433  
Qy 1146 yrGlnGlyAsnProLeuArgGlyThrCysTyraTyraThrLeuLeuIleAspTyraGlnPheT 1166  
Db 3434 ACCAAGGAAACCTCTCAGAGGAACATGTTATTATATCTTCTTATTGACTATCAGTTTCA 3493  
Qy 1166 hrPheSerLeuSerGlnGluAspAspArgTyraTyraThrAlaIleAsnPheValAlaThrP 1186  
Db 3494 CCTTTTAGTCTATCCCGAGAGATGATCGCTATTACACAGCTATCAATTTTGGGTACTC 3553  
Qy 1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206  
Db 3554 CTGACGAAACAAACAGGATTTGGACATGTTTCATCAATGCCCTCCAAAGAAATTTCAACCTCA 3613  
Qy 1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluMetProValV 1226  
Db 3614 ACATCACCTGGCTGCCAGTTCTCAGCTGGAAACCCAGGCTGGAGAAAGATGCCCTGTTG 3673  
Qy 1226 alSerLysThrAsnIleLysGluTyraLysAspSerPheSerAsnGluLysPheAspPheA 1246  
Db 3674 TTTCAAAACCAACATTAAGGAGTACAAGATAGTTTCTCTAATGAGAAGTTTGATTTTC 3733  
Qy 1246 rgAsnHisProAsnIleThrPhePheValTyraValSerAsnPheThrTrpProIleLysI 1266  
Db 3734 GCAACCCACCAATATCACTTTCTTGTATTGTATGTAGTAAATTTCACTTGGCCCAATCAAAA 3793  
Qy 1266 leGlnIleAlaPheSerGlnHisSerAsnPheMetAspLeuValGlnPhePheValThrP 1286  
Db 3794 TTCAGATTGCCCTTCTCTCAGCACACGCAATTTTATGGACCTGGTACAGTTCTTCGTGACTT 3853



486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyGlyTyrGlyHisSerSerV 506  
1454 CATGGAGTATTATACACACCCAGGGTCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513  
506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrIysAlaPheSerAlaA 526  
1514 TTTACGACCATAGGACAGGGCCCTATACGTTTCATGGTGCTACAGGGCTTTTCAGTGCCA 1573  
526 snIysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpTrpI 546  
1574 ATAAGTACCGGCTTGAGATGATCTCTACCGATATGATGTGATACCCAGATGTGACCA 1633  
546 leuLeuIysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566  
1634 TTTCTTAAGGACACCCGATTTTCCGTTACTTGCACACAGCTGTGATAGTGAAGCA 1693  
566 etLeuValPheGlyGlyValAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586  
1694 TGCTGGTGTGTTGGGGGAAACACACACACATGATCTATGACCATGGCGGCGAATGCT 1753  
586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgP 606  
1754 TCTCTTCAGATTTTCATGGCTATGACATTCCTGTGACCGCTGCTCAGTGCCTTCCAGAC 1813  
606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626  
1814 CTGATCTCCACCAATGATGTCACAGATTTGGCCATTCAGAGCTTTACACACAGACCA 1873  
626 etTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerG 646  
1874 TGTATGTGTTGGTGGTTCATAGTCTCTCTCAGCGACATCTCTGGTATTCACTCGG 1933  
646 luClnCysAspAlaHisArgSerGluAlaAlaCysLeuAlaAlaGlyProGlyIleArgC 666  
1934 AACAGTGTATGGCATCGAGTGAAGCCGCTGTTTACAGCAGGACCTGGTATTTCGGT 1993  
666 ysValTrpAsnThrGlySerSerGlnCysIleSerTrpAlaLeuAlaThrAspGluGlnG 686  
1994 GTGTGTGGAAACACAGGGTGTCTCAGTGTATCTGTGGGCGTGGCACTGATGAACAG 2053  
686 luGluIysLeuIysSerGluCysPheSerIysArgThrLeuAspHisAspArgCysAspG 706  
2054 AAGAAAAGTTAAATCAGAATGTTTTCAAAAAGAACTCTTGACCATGACAGATGTGACC 2113  
706 luHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTrpCysAsnAspH 726  
2114 AGCACACAGATTGTTACAGCTGCACAGCCCAACCAATGACTGCCACTGGTCAATGACC 2173  
726 isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746  
2174 ATTGTGTCCCGAAGAACACACAGCTGCTCAGAGGCCAGATCTCCATTTTATGATAGA 2233  
746 snCysProIysAspAsnProMetTyrTyrCysAsnIysIysThrSerCysArgSerCysA 766  
2234 ATTGCCCCCAAGGATAACCTATGCTACTGTAAACAAGAACCAAGCTCAGAGAGCTGTG 2293  
766 laLeuAspGlnAsnCysGlnTrpGluProArgAsnGlnGluCysIleAlaLeuProGluA 786  
2294 CCCTGGACCAAGACTCCAGTGGAGGCCCGGNAATCAGAGTGCATTTGCCCTGCCGAAA 2353  
786 snIleCysGlyIleGlyTrpHisLeuValGlyAsnSerCysLeuIysIleThrAlaAl 806  
2354 ATATCTGTGCATTTGGCTGGCAATTTGGTTGGAACTCATGTTTGAATAATTAATCTGCA 2413  
806 ysGluAsnTyrAspAsnAlaIysLeuPheCysArgAsnHisAsnAlaLeuAlaSerL 826  
2414 AGGAGAATTATGCAATGCTAAATTTGTTGTAGGAACCAATGGCCCTTTTGGCTTCTC 2473  
826 euThrThrGlnIysIysValGluPheValLeuIysGlnLeuArgIleMetGlnSerSerG 846  
2474 TTACACCCAGAGAGGTAGAAATTTGTTCTTAAGCAGCTGCGAATAATGCAATCATCTC 2533

846 lnSerMetSerLysLeuThrLeuThrProTrpValGlyLeuArgIysIleAsnValSerT 866  
2534 AGAGCATGTCCAAGCTACCTTAACCCCATGGTTCGGCTTCGGAAGATCAATGTGTCT 2593  
866 yTrpCysTrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerG 886  
2594 ACTGTGTCTGGAGAGATATGTCCTCCCATTTTACAATAGTATTACTACAGTGGATGGCTGT 2653  
886 luProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuIysA 906  
2654 AGCCAGATGATGCTGGATTCTGTGGAAATTTTATCAGAACCCAGTACTCGGAGACTGAAG 2713  
906 laAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgProAlaAsnHisSerA 926  
2714 CTGCAACCTGATCAACCCACTCAATGGTAGTGTCTGTGAAGGCCCTGCAAAACACAGTG 2773  
926 laLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAspCysThrSerGlyS 946  
2774 CTAGCAGTGGCGGACACCATGTCCTTGAGGACAGCATGTGGAGATTGCACACAGCGCA 2833  
946 erSerGluCysMetTrpCysSerAsnMetIysGlnCysValAspSerAsnAlaTyrValA 966  
2834 GCTCTGAGTGCATGTGGTGACCAACATGAAGCAGTGTGTGACTCCAAATGCCTATGTGG 2893  
966 laSerPhePheGlyGlnCysMetGluTrpTyrThrMetSerThrCysProProGluA 986  
2894 CCTCTCTCCCTTTTGGCCAGTGTATGGAATGGTATACATGAGCACCTGCCCCCTGAAA 2953  
986 snCysSerGlyTyrCysThrCysSerHisCysLeuGluGlnProGlyCysGlyTrpCysT 1006  
2954 ATTGTTCAGGCTACTGTACCTGTAGTCAATGCTGGAGCAACACAGGCTGTGGCTGGTGA 3013  
1006 hrAspProSerAsnThrGlyIysGlyIysCysIleGluGlySerTyrIysGlyProValL 1026  
3014 CTGATCCAGCAATACTGGCAAGGAAATGCATAGAGGGTTCCTATAAAGAGCACAGTGA 3073  
1026 ysMetProSerGlnAlaProThrGlyAsnPheTyrProGlnProLeuLeuAsnSerSerM 1046  
3074 AGATGCCTTCCAGAGCCCTACAGAAATTTCTATCCACAGCCCTGCTCAATTCACAGCA 3133  
1046 etCysLeuGluAspSerArgTyrAsnTrpSerPheIleHisCysProAlaCysGlnCysA 1066  
3134 TGTGTCTAGAGGACAGACAGATACAACTGCTCTTTCATTCATCTGCTCCAGCTTGGCAATGCA 3193  
1066 snGlyHisSerLysCysIleAsnGlnSerIleCysGluLysCysGluAsnLeuThrArg 1086  
3194 AGGCCACAGTAAATGATCAATCAGAGCATCTCTGAGAAGTGTGAGAACCTGACCAACAG 3253  
1086 lyLysHisCysGluThrCysIleSerGlyPheTyrGlyAspProThrAsnGlyGlyLysC 1106  
3254 CCAAGCATGGAGACCTGCATATCTGGCTTCTACGGTGTATCCCAATGGAGGGAAT 3313  
1106 ysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAsnThrAsnThrGlyLysCysP 1126  
3314 GTCAGCATGCAAGTGAATGGGCACCGCTCTCTGTGCAACACACCAACCGGGCAAGTGTCT 3373  
1126 heCysThrThrLysGlyValIysGlyAspGluCysGlnLeuCysGluValGluAsnArgT 1146  
3374 TCTGACCAACCAAGGGGTCAAGGGGAGCGAGTCCAGCATGTGTGAGGTAGAAATCGAT 3433  
1146 yTrpGlnGlyAsnProLeuArgGlyThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheT 1166  
3434 ACCAAGGAAACCCCTCTCAGAGGAAATGTTATTATATCTCTTCTTATTGACTATCAGTTCA 3493  
1166 hrPheSerLeuSerGlnGluAspArgTyrTyrThrAlaIleAsnPheValAlaThrP 1186  
3494 CCTTATGCTATCCAGAGAGATGATCGTATTTACAGACGATATCAATTTTGTGCTACTC 3553  
1186 roAspGluGlnAsnArgAspLeuAspMetPheIleAsnAlaSerLysAsnPheAsnLeuA 1206  
3554 CTGACGAAACAAACAGGATTTGGACATGTTTCATCAATGCTCCCAAGAAATTTCAACCTCA 3613  
1206 snIleThrTrpAlaAlaSerPheSerAlaGlyThrGlnAlaGlyGluGluMetProValV 1226

```
Db 3614 ACATCACTGGCTGCGAGTTTCTCAGCTGAACCCAGGCTGAGAGAGATGCTGTGTTG 3673
Qy 1226 alSerLysThrAsnIleLysGluTyrLysAspSerPheSerAsnGluLysPheAspPheA 1246
Db 3674 TTTCAAAACCAACATTAAAGAGTACAAAGATAGTTCTCTAATGAGAAGTTTGAATTTTC 3733
Qy 1246 xAsnHisProAsnIleThrPhePheValTyrValSerAsnPheThrTTPProIleLysI 1266
Db 3734 GCAACACCCAAATATACATTTCTTTTATGTAGTAATTTCACTGGCCCATCAAAA 3793
Qy 1266 leGln 1267
Db 3794 TTCAG 3798

RESULT 3
US-09-245-041-12
; Sequence 12, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Naegle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 12
; LENGTH: 6370
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-245-041-12

Alignment Scores:
Pred. No.: 0 Length: 6370
Score: 503.00 Matches: 516
Percent Similarity: 99.81% Conservative: 0
Best Local Similarity: 99.81% Mismatches: 0
Query Match: 35.20% Indels: 1
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-245-041-12 (1-6370)
Qy 914 AsnGlySerValCysGluArgProAlaAsnHisSerAlaLys-GlnCysArgThrProCy 933
Db 506 AATGGTAGTGTCTGTGAAGCCCTGCAACACCAAGTCTAAGCAGTGGCCGACACATG 565
Qy 933 sAlaLeuArgThrAlaCysGlyAspCysThrSerGlySerSerGluCysMetTrpCysSe 953
Db 566 TGCCTTGAGGACAGCATGTGGAGATTGACACGCGGAGCTCTGAGTGATGTGGTGAC 625
Qy 953 rAsnMetLysGlnCysValAspSerAsnAlaTyrValAlaSerPheProPheGlyGlnCy 973
Db 626 CAACATGAAGCAGTGTGTGACTCAATGCTATGTGGCTCTCTCTCTCTCTCTCTCTCTCT 695
Qy 973 sMetGluTyrThrMetSerThrCysProProGluAsnCysSerGlyTyrCysThrCy 993
Db 686 TATGGAATGTATACGATAGCAGCTGCCCCCTGAAATTTGTTCAGGCTACTGTACCTG 745
Qy 993 sSerHisCysLeuGlnProGlyCysGlyTyrCysThrAspProSerAsnThrGlyLy 1013
Db 746 TAGTCATTGCTTGGAGCAACAGGCTGTGGTGTGTACTGTATCCCAAGCAATAGTGCAA 805
Qy 1013 sGlyLysCysIleGluGlySerTyrLysGlyProValLysMetProSerGlnAlaProTh 1033
Db 806 AGGAAATGTCATAGAGGTTCTTATTAAGGACCAAGTGAAGTGCCTTCGCAAGCCCTAC 865
```

```
Qy 1033 rGlyAsnPheTyrProGlnProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTy 1053
Db 866 AGGAAATTTTCTATCCACAGGCCCTGTCTCAATTCAGCATGTGTCTAGGACAGCAGATA 925
Qy 1053 rAsnTrpSerPheIleHisCysProAlaCysGlnCysAsnGlyHisSerLysCysIleAs 1073
Db 926 CAACCTGGTCTTTCATTCACTGTCCAGCTTGCATGCAAGCGCCACAGTAAATGCATCAA 985
Qy 1073 nGlnSerIleCysGluLysCysGluAsnLeuThrThrGlyLysHisCysGluThrCysI 1093
Db 986 TCAGAGCATCTGTGAGAAGTGTGAGAACCTGACCCACAGGAAGCACTGCGAGACCTGCAT 1045
Qy 1093 eSerGlyPheTyrGlyAspProThrAsnGlyLysCysGlnProCysLysCysAsnG 1113
Db 1046 ATCTGGCTTCTACGGTGTATCCCAATGGAGGAGAAATGTGAGCCATGCAAGTGCATAGG 1105
Qy 1113 yHisAlaSerLeuCysAsnThrAsnThrGlyLysCysPheCysThrThrLysGlyValLy 1133
Db 1106 GCACGGTCTCTGTGCAACCAACACGCGCAGTGTCTTGACACCAACCAAGGGCGTCAA 1165
Qy 1133 sGlyAspGluCysGlnLeuCysGluValGluAsnArgTyrGlnGlyAsnProLeuArgG 1153
Db 1166 GGGGACGAGTCCAGCTATGTGAGGTAGAAATCGATACCAAGAAACCCCTCTCAGAGG 1225
Qy 1153 yThrCysTyrTyrThrLeuLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAs 1173
Db 1226 AACATGTTTATATACTCTCTTATGACTATCAGTTACCTTTAGTCTTATCCAGGAGAA 1285
Qy 1173 pAspArgTyrTyrThrAlaIleAsnPheValAlaThrProAspGluGlnAsnArgAspLe 1193
Db 1286 TGATCGCTATTACACAGCTATCAATTTTGTGGCTACTCTCTGACGAAACAAACAGGATTT 1345
Qy 1193 uAspMetPheIleAsnAlaSerLysAsnPheAsnLeuAsnIleThrTrpAlaAlaSerPh 1213
Db 1346 GGACATGTTTCATCAATGCCTCCAAAGATTTCAACCTCAACATCACCTGGGCTGCCAGTTT 1405
Qy 1213 eSerAlaGlyThrGlnAlaGlyGluGluMetProValValSerLysThrAsnIleLysG 1233
Db 1406 CTCAGCTGAACCCAGGCTGGAGAGAGATGCTGTGTTTCAAAACCAACATTAAGGA 1465
Qy 1233 uTyrLysAspSerPheSerAsnGluLysPheAspPheArgAsnHisProAsnIleThrPh 1253
Db 1466 GTACAAAGATAGTTTCTCTAATGAGAAGTTTGATTTTCGCAACCCACCAATATACATTT 1525
Qy 1253 ePheValTyrValSerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHi 1273
Db 1526 CTTTGTATTATGTCAATAATTTCACTGGCCCATCAAAATTCAGATTGCTTCTCTCAGCA 1585
Qy 1273 sSerAsnPheMetAspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLe 1293
Db 1586 CAGCAATTTTATGGACCTGTGTACAGTTCTCTCGTCACTTCTTCAGTTGTTTCTCTCTCT 1645
Qy 1293 uLeuLeuValAlaAlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgAr 1313
Db 1646 GCTCTGTGTGGCTGTGTGGTTTGAAGATCAAAAGTTTGTGGGCTTCAGACGCTAG 1705
Qy 1313 sGluGlnLeuLeuArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnVa 1333
Db 1706 AGAGCAACTTCTTCAGAGATGCAACAGATGCGCAGCGCTCCCTTTGCTCTGTAATGT 1765
Qy 1333 lAlaLeuGluThrAspGluGluProProAspIleLysGlySerIleLysThrValPr 1353
Db 1766 GCGCTTGGAAACAGATGAGGAGCCCTCCGTGATCTTATTTGGGGGAGTATAAAGACTGTTC 1825
Qy 1353 oLysProIleAlaLeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheVa 1373
Db 1826 CAAACCCATTCGACTGGAGCCGTGTTTGGCAACAAAGCCGCTGCTCTCTCTCTGTTGT 1885
Qy 1373 lArgLeuProArgGlyLeuGlyIleProProProGlyGlnSerGlyLeuAlaValAl 1393
Db 1886 GAGGCTCCCTCGAGGCCCTGGGTGGCATCCCTCCCTGGGCGAGTCAGGCTGTGTGTGGC 1945
Qy 1393 aSerAlaLeuValAspIleSerGlnGlnMetProIleValTyrLysGluLysSerGlyAl 1413
```

Db 1946 CAGCGCCCTGGTGGACATTTCTCAGCAGATGCCGATAGTGTACAGGAGAGTCCAGAGC 2005  
Qy 1413 aValArgAsnArglyslnglnProProAlaGlnProGlyThrCysile 1429  
Db 2006 CGTGAGAAACCGAGAGCAGAGCCCTGCAGAGCTGGACCTGCATC 2054

## RESULT 4

US-09-245-041-18  
; Sequence 18, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagie, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; CURRENT FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 18  
; LENGTH: 2625  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-245-041-18

## Alignment Scores:

Pred. No.: 0 Length: 2625  
Score: 496.00 Matches: 696  
Percent Similarity: 99.43% Conservative: 0  
Best Local Similarity: 99.43% Mismatches: 2  
Query Match: 34.71% Indels: 4  
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-245-041-18 (1-2625)

Qy 87 AlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106  
Db 256 GCGGCGCGGTGTCGCGCTCGCGCGAGCGAGCCCAAGAAATGTACCGCGCGTGTGC 315  
Qy 107 AsnGlyClyArgCysAsnProGlyThrGlnCysValCysProAlaGlyTrpValGly 126  
Db 316 AACGGCGTCCCTGCAACCTGGCACCGCCAGTGGCTGTGCCCCCGCTGGGTGGC 375  
Qy 127 GluGlnCysGlnHisCysGlyArgPheArgLeuThrGlySerSerGlyPheValThr 146  
Db 376 GAGCAATGCCAGCACTGCGGGGCGCGCTTCAGACTAACTGGATCTTCTGGGTTGTGACA 435  
Qy 147 AspGlyProGlyAsnTrpLysTrpLysThrLysCysThrTrpLeuIleGluGlnPro 166  
Db 436 GATGGACCTGGAATATATAATCAAAAGCAAGTGCACGTGGCTCATTTGAAGACAGCCA 495  
Qy 167 AsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHisLeu 186  
Db 496 AATAGAATAATGAGACTTCGTTTCATCATTTTGTCTACAGAGTGTAGTTGGACCATTTA 555  
Qy 187 TyrValTrpAspGlyAspSerIleTyrAlaProLeuValAlaAlaPheSerGlyLeuIle 206  
Db 556 TATGTTTATGATGGGACCTCAATTTATGCAACCGCTAGTTGCTGCATTTAGTGGCTCAT 615  
Qy 207 ValProGluArgAspGlyAsnGluThrValProGluValValAlaThrSerGlyTyrAla 226  
Db 616 GTTCTCGAGAGATGGCAATGAGACTGTCTCCCTGAGGTTGTTGCCACATCAGGTTATGCC 675  
Qy 227 LeuLeuHisPhePheSerAspAlaAlaTyrAsnLeuThrGlyPheAsnIleThrTrpSer 246  
Db 676 TTGCTGCATTTTATTAGTGATGCTGCTTATAATTTGACTGGATTTAATATTACTTACACT 735

Qy 247 PheAspMetCysProAsnAsnCysSerGlyArgGlyGluCysLysIleSerAsnSerSer 266  
Db 736 TTTGATATGTGTCCAAATAAATCTGCTCAGCGCAGGAGAGTGTAAAGATCAGTAATAGCAGC 795  
Qy 267 Glu-ThrValGluCysGluCysSerGluAsnTrpLysGlyGluAlaCysAspIleProHi 286  
Db 796 GA-THACTGTGAATGTGAATGTTCTGAAACTGGAAGGTGAAGCATGTGACATTCCTCA 854  
Qy 286 sCysThrAspAsnCysGlyPheProHisArgGlyIleCysAsnSerSerAspValArgG1 306  
Db 855 CTGTACAGACAACCTGTGCTTTTCTCATCAGGAGCATCTGCAATCAAGTGTGTCAGAGG 914  
Qy 306 yCysSerCysPheSerAspTrpGlnGlyProGlyCysSerValProValProAlaAsnG1 326  
Db 915 ATGCTCCTGCTTCTCAGACTGGCAGGTCCTGGATGTTTCAGTTCCTGTACAGTAACCA 974  
Qy 326 nSerPheTrpThrArgGluGluTyrSerAsnLeuLysLeuProArgAlaSerHisLysAl 346  
Db 975 GTCATTTTGGACTCGAGAGGAATATTCTAACTTAAAGCTCCCGAGAGCATCTCATAAAGC 1034  
Qy 346 aValValAsnGlyAsnIleMetTrpValValGlyGlyTyrMetPheAsnHisSerAspTy 366  
Db 1035 TGTGTCATATGGAACAATATATGTGGTGTGTGGAGGATATATGTTCACCACTCAGATTA 1094  
Qy 366 rAsnMetValLeuAlaTyrAspLeuAlaSerArgGluTrpLeuProLeuAsnArgSerVa 386  
Db 1095 TAACATGTTCTAGCGTATGACCTTGTCTTAGGGAGTGGCTTCACATAACCGTCTGT 1154  
Qy 386 lAsnAsnValValValArgTyrGlyHisSerLeuAlaLeuTyrLysAspLysIleTyrMe 406  
Db 1155 GAACAATGTGTTGTAGATATGTCATTTCTTGGCATTATACAAGGATAAAATTTACAT 1214  
Qy 406 tTyrGlyGlyLysIleAspPro-ThrGlyAsnValThrAsnGluLeuArgValPheHisI 426  
Db 1215 GTATGGAGAAATTTGATTC-AACTGGAAATGACCAATGAGTGTAGAGTGTTCACA 1273  
Qy 426 leHisAsnGluSerTrpValLeuLeuThrProLysAlaLysGluGlnTyrAlaValValG 446  
Db 1274 TTCATAATGAGTGCATGGGTGTGTTGACCCCTAAGGCAAGAGCAGTATCGAGTGTG 1333  
Qy 446 yHisSerAlaHisIleValThrLeuLysAsnGlyArgValValMetLeuValIlePheG 466  
Db 1334 GGCACCTCTGCACACATTTGTACACTGAAGAATGCCGAGTGGTTCATGCTGCTATCTTTG 1393  
Qy 466 yHisCysProLeuTyrGlyTyrIleSerAsnValGlnGluTyrAspLeuAspLysAsnT 486  
Db 1394 GTCACTGCCCTCTCTATGATATATAGCAATGTGCAGAAATATGATTTGGATAGAACA 1453  
Qy 486 hrTrpSerIleLeuHisThrGlnGlyAlaLeuValGlnGlyTyrGlyHisSerServ 506  
Db 1454 CATGAGTATATTACACACCCAGGCTGCCCTTGTGCAAGGGGGTTACGGCCATAGCAGTG 1513  
Qy 506 alTyrAspHisArgThrArgAlaLeuTyrValHisGlyGlyTyrLysIlePheSerAlaA 526  
Db 1514 TTTAGACCATAGGACCCAGGGCCCTTACGTTTCATGTTGCTGAGGCTTTCAGTGGCA 1573  
Qy 526 snLysTyrArgLeuAlaAspAspLeuTyrArgTyrAspValAspThrGlnMetTrpThrI 546  
Db 1574 ATAAGTACCGGCTTGCAGATGATCTTACCAGATATGATGATGATGATGATGATGATGATG 1633  
Qy 546 leLeuLysAspSerArgPhePheArgTyrLeuHisThrAlaValIleValSerGlyThrM 566  
Db 1634 TTTCTTAAGACAGCGCGATTTTTCGCTTACTTGCACACAGCTGTGATAGTGTAGTGAACCA 1693  
Qy 566 etLeuValPheGlyClyAsnThrHisAsnAspThrSerMetSerHisGlyAlaLysCysP 586  
Db 1694 TGCTGGTGTGTTGGGGGAAACACACATGATGACATCTATGAGCGATGGCCCAATGCT 1753  
Qy 586 heSerSerAspPheMetAlaTyrAspIleAlaCysAspArgTrpSerValLeuProArgp 606  
Db 1754 TCTCTTCAGATTTTCATGCGCTTATGACATTTGCTGACCGCTGTGTCAGTGTCTCCAGAC 1813  
Qy 606 roAspLeuHisHisAspValAsnArgPheGlyHisSerAlaValLeuHisAsnSerThrM 626



```

Db      1814  CTGATCTCCCAATGATGTCACAGATGGCCATTCAGCAGTCTTACACACAGACCA 1873
Qy      626  etTyrValPheGlyGlyPheAsnSerLeuLeuSerAspIleLeuValPheThrSerG 646
Db      1874  TGTATGTTTCCGTTGGTTTCAATAGTCTCTCTCCTCAGCAGACATCTCTGGTATTCACTCG 1933
Qy      646  luGlnCysAspAlaHisArgSerGluAlaCysLeuAlaGlyProGlyIleArgC 666
Db      1934  AACAGTGTGATGGCATCGGAGTGAAGCCCTGTGTAGCAGCAGACCTGGTATTTCGT 1993
Qy      666  ysValTyrAsnThrGlySerSerGlnCysIleSerTyrAlaLeuAlaThrAspGluGlnG 686
Db      1994  GTGTGTGGAACACAGGCTCTCTCAGTGTATCTCGTGGGCGCTGGCACTGATGAACAAG 2053
Qy      686  luGluLysLeuLysSerGluCysPheSerIysArgThrLeuAspHisAspArgCysAspG 706
Db      2054  AAGAAAGATTAAATCAAGATGTTTTTCCAAAGAACTCTTGACCATGACAGATGTGACC 2113
Qy      706  lnHisThrAspCysTyrSerCysThrAlaAsnThrAsnAspCysHisTyrCysAsnAspH 726
Db      2114  AGACACAGATTGTACAGCTGCACAGCCACACCAATGACTGCCACTGGTGCATGACC 2173
Qy      726  isCysValProArgAsnHisSerCysSerGluGlyGlnIleSerIlePheArgTyrGluA 746
Db      2174  ATTGTGTCCCGAGAACACACAGCTGCTCAGAGGCCAGATCTCCATTTTTAGGTATGAGA 2233
Qy      746  snCysProLysAspAsnProMetTyrTyrCysAsnLysLysThrSerCysArgSerCysA 766
Db      2234  ATTGCCCCAGGATAACCCCTATTGACTACTGTAACTGTAACAGAGACCACTGCAGGAGCTGTG 2293
Qy      766  laLeuAspGlnAsnCysGlnTyrGluProArgAsnGlnGluCysIleAlaLeuPro 784
Db      2294  CCCTGGACCAAGACTGCCAGTGGAGCCCGGAATCAGGAGTGCATTCGCTGCC 2349

```

## RESULT 5

```

US-09-245-041-1
; Sequence 1, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8827
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-1

```

```

Alignment Scores:
Pred. No.:      3.2e-182      Length:      8827
Score:          206.00      Matches:      576
Percent Similarity: 96.00%      Conservative: 0
Best Local Similarity: 96.00%      Mismatches: 12
Query Match:      14.42%      Indels:      24
DB:              3          Gaps:      0

```

US-09-787-097-12 (1-1429) x US-09-245-041-1 (1-8827)

```

Qy      842  MetGlnSerGlnSerMetSerLysLeuThrLeuThrProTyrValGlyLeuArgLys 861
Db      2790  ATGCAATCATCTCAAGATATGTCACAGCTCACTGACTCCATCGGTGGTCTTCGGAAG 2849

```

```

Qy      862  ileAsnValSerTyrTyrCysTyrGluAspMetSerProPheThrAsnSerLeuLeuGln 881
Db      2850  ATCAATGTGTCTTACTGGTCTGGAGGATATGTCTCCATTCAAAATAGTTTGTCTCAG 2909
Qy      882  TrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGluProSerThr 901
Db      2910  TGGATGCCATCTGAGCCACAGTATGCTGGCTTCTGTGGGATCTTGTGAGAGCCTAGTACT 2969
Qy      902  ArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCysGluArgPro 921
Db      2970  CGGGATTAAGGTGTCGAACCTGCATCAACCTCTCAATGGCAGCTCTCTGAAAGCCCT 3029
Qy      922  AlaAsnHisSerAlaLysGlnCysArgThrProCysAlaLeuArgThrAlaCysGlyAsp 941
Db      3030  GCAAAACACAGTGCAGAGCAGTGCAGACACCATGTGCCCTGGCGACAGCGTGTGGCGA - 3088
Qy      942  -CysThrSerGly-SerSerGluCysMetTyrCysSerAsnMetLysGlnCysValAsps 961
Db      3089  GTGCACACTAGCAG - CAGCTCGGAGTGCATGTGTGTCAGTACATGACATGACAGTGTGGACT 3147
Qy      961  erAsnAlaTyrValAlaSerPheProPheGlyGlnCysMetGluTyrTyrThrMetSerT 981
Db      3148  CCAATGCCCTAGTGCCTCTCTTCCCTTTTGGCCAGTGTATGGAATGTATACGATGACGA 3207
Qy      981  hr-CysProProGluAsnCysSerGlyTyrCysThrCysSerHisCysLeuGlnPro 1000
Db      3208  G-CTGCCACACTGAAATTTGCTCTGGCTACTGTATACCTGCAGCCCAATGCTTGGAGCAGCA 3266
Qy      1001  GlyCysGlyTyrCysThrAspProSerAsnThrGlyLysGlyLysCysIleGluGlySer 1020
Db      3267  GGCTGTGGTGGTGTACTGTATCTCTAGCAATACTGGGAAGGAAATGTATTAGGGCAGC 3326
Qy      1021  TyrLysGlyProValLysMetProSerGlnAlaProThr--GlyAsnPhe-TyrProGln 1039
Db      3327  TATAAAGGACCTGTGAAGATGCCCTGCAGGCCCTC-TG-CAGGAAATGT-GTATCCACAG 3383
Qy      1040  ProLeuLeuAsnSerSerMetCysLeuGluAspSerArgTyrAsnTyrSerPheIleHis 1059
Db      3384  CCCCTTCTGAACTCAGCATGTGTCTAGAGGACAGCAGATACACTGGTCTTTCATTAC 3443
Qy      1060  CysProAlaCysGlnCysAsnGlyHisSerLysCysIleAsnGlnSerIleCysGluLys 1079
Db      3444  TGTCCAGCTTGCCAGTGCACCGACACAGCAATGCATCAACAGAGATATCTGTGAGAAG 3503
Qy      1080  CysGluAsn-LeuThrThrGlyLysHisCysGluThrCysIleSerGlyPheTyrGlyAs 1099
Db      3504  TGTGAGGA-CTGTACACCGGCAAGCAGCTGCAGACCTGCATATCTGGCTTCTATGTGA 3562
Qy      1099  pProThrAsnGlyGlyLysCysGlnProCysLysCysAsnGlyHisAlaSerLeuCysAs 1119
Db      3563  CCCGACTAATGAGGCAAAATGTACAGCATGCAAGTGCATGGGCACGCATCACTGTGCA 3622
Qy      1119  mThrAsnThrGlyLysCysPheCysThrThrLysGlyValLysGlyAspGluCysGlnLe 1139
Db      3623  CACCAACACCGGCAAGTGTCTGTACCCAAAGGTGTCAAGGGGACGAGTGCAGCT 3682
Qy      1139  uCysGluValGluAsnArgTyrGlnGlyAsnProLeuArg-GlyThrCysTyrTyrThrL 1159
Db      3683  ATGTGAGGTAGAAATCGATACCAAGGAACCCCTCTCAA-AGGAACATGCTACTATACCC 3741
Qy      1159  euLeuIleAspTyrGlnPheThrPheSerLeuSerGlnGluAspAspArgTyrTyrThrA 1179
Db      3742  TTCTCATTTAGCTATCAGTTTACCTTTAGCCTGTCTCCAGGAGACGCGCTACTACACAG 3801
Qy      1179  laIleAsnPheValAlaThrProAspGluGlnAsnArgAspLeuAspMetPheIleAsnA 1199
Db      3802  CCATCACTTTGTGGCTACTCTCTGATGACAAAACAGGGATTTGACATGTTTCATCAATG 3861
Qy      1199  laSerLysAsnPheAsnLeuAsnIleThrTyrAlaAla-SerPheSer-AlaGlyThrGln 1218
Db      3862  CTTCCAAAATCTCAACCTCAACATCACTGGGCCAC-CAGCTTCCC-AGCCGGAACCCCA 3919
Qy      1218  nAla-GlyGluGluMet-ProValValSerLysThrAsnIleLysGluTyrLysAspSer 1237

```

Db 3920 GAC TGGAGAGAGGT-GCTGTGTGTTTCAAAACCAACATCAAGGATCAACAGATACG 3977  
Qy 1238 PheSerAsnGlnLysPheAspPheArgAsnHisProAsnIleThrPhePheValTyrVal 1257  
Db 3978 TTCTCTAATGAGAAATTTGATTTTCGCAACATCCAAACATCACTTTCTTTGTTATGTC 4037  
Qy 1258 SerAsnPheThrTrpProIleLysIleGlnIleAlaPheSerGlnHisSerAsnPheMet 1277  
Db 4038 AGTAATTTCACTTGGCCCATCAAAATTCAGATTGCTTCTCCAGCAGACCACTTCATG 4097  
Qy 1278 AspLeuValGlnPhePheValThrPhePheSerCysPheLeuSerLeuLeuValAla 1297  
Db 4098 GACCTGTCAGATTCTTCGTCAGCTTTCTTCAGTTGTTTCTCTCGCTGCTTCTGTTGGT 4157  
Qy 1298 AlaValValTrpLysIleLysGlnSerCysTrpAlaSerArgArgGlnGlnLeuLeu 1317  
Db 4158 GCAGTGGTCTGGAGATCAAGCAGAGCTGTTGGGCATCCAGGGGAGAGAGCAACTTCT 4217  
Qy 1318 ArgGluMetGlnGlnMetAlaSerArgProPheAlaSerValAsnValAlaLeuGluThr 1337  
Db 4218 CGGAGATGCAACAGATGCCAGCCGCCCTTTGCTTCTGTAACGTTGCCTTGGAACA 4277  
Qy 1338 AspGluGluProProAspIleGlyGlySerIleLysThrValProLysProIleAla 1357  
Db 4278 GATGAGAGGCTCTGATCTTATTTGGGGGAGTATAAGACTGTTCCCAACCCCATGCA 4337  
Qy 1358 LeuGluProCysPheGlyAsnLysAlaAlaValLeuSerValPheValArgLeuProAsp 1377  
Db 4338 CTGAGCCGTTGTTGGCAACAAAGCGCTGCTCTCTGTGTTGTGAGGCTCCCTCGA 4397  
Qy 1378 GlyLeuGlyGlyLeuProProGlyGlnSerGlyLeuAlaValAlaSerAlaLeuVal 1397  
Db 4398 GGCTCGGTGGCATCCCTCTCTCGGCGAGTCAGGCTTGTGTTGGCCAGCGCCCTGGTG 4457  
Qy 1398 AspIleSerGlnMetProIleValTyrLysGluLysSerGlyValAlaValArgAsnArg 1417  
Db 4458 GACATTTCTCAGCAGATGCCATAGTGTACAGAGAGAAAGTCAGGAGCCGTGGAACCCG 4517  
Qy 1418 LysGlnGlnProAlaGlnProGlyThrCysIle 1429  
Db 4518 AAGCAGCAGCCCTCGCACAGCTGGGACCTGCATC 4553

## RESULT 6

US-09-245-041-8  
; Sequence 8, Application US/09245041  
; Patent No. 6274339  
; GENERAL INFORMATION:  
; APPLICANT: Moore, K.  
; APPLICANT: Nagle, D.  
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT  
; OF BODY WEIGHT DISORDERS INCLUDING OBESITY  
; FILE REFERENCE: 7853-136  
; CURRENT APPLICATION NUMBER: US/09/245,041  
; EARLIER FILING DATE: 1999-02-05  
; EARLIER APPLICATION NUMBER: 60/093,630  
; EARLIER FILING DATE: 1998-07-21  
; EARLIER APPLICATION NUMBER: 60/104,978  
; EARLIER FILING DATE: 1998-10-20  
; NUMBER OF SEQ ID NOS: 131  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 8  
; LENGTH: 2419  
; TYPE: DNA  
; ORGANISM: Mus musculus

US-09-245-041-8  
Alignment Scores:  
Pred. No.: 5,35e-78 Length: 2419  
Score: 94.00 Matches: 94  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 6.58% Indels: 0

DB: 3 Gaps: 0  
US-09-787-097-12 (1-1429) x US-09-245-041-8 (1-2419)  
Qy 513 AlaLeuTyrValHisGlyGlyTyrLysAlaPheSerAlaAsnLysTyrArgLeuAlaAsp 532  
Db 1803 GCTCTGTACGTTTCATGGTGGCTTACAGGCTTTCAGCGCAACAAATACCGCTTTCAGAT 1862  
Qy 533 AspLeuTyrArgTyrAspValAspThrGlnMetTrpThrIleLeuLysAspSerArgPhe 552  
Db 1863 GACCTCTACATACGATGATGGATCTCAGATGTGGACCATCTTAAGGACAGCGGATTT 1922  
Qy 553 PheArgTyrLeuHisThrAlaValIleValSerGlyThrMetLeuValPheGlyAsn 572  
Db 1923 TTCCGTTACTTGCATACAGCTGTATAGTAGTGGAAACCATGCTGGTGTGTTGGAGGAAC 1982  
Qy 573 ThrHisAsnAspThrSerMetSerHisGlyAlaLysCysPheSerSerAspPheMetAla 592  
Db 1983 ACACAATGACACTTCCATGAGCCAGCGTGCCTCTCTCTCGGACTTCATGGCT 2042  
Qy 593 TyrAspIleAlaCysAspArgTyrSerValLeuProArgPro 606  
Db 2043 TATGACATTTGTTGTGACCGATGTCAGTGTCTCCAGACCT 2084

## RESULT 7

US-09-702-705-966/c  
; Sequence 966, Application US/09702705  
; Patent No. 6504010  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Tongtong  
; APPLICANT: Bangur, Chaitanya S.  
; APPLICANT: Lodes, Michael A.  
; APPLICANT: Fanger, Gary  
; APPLICANT: Vedwick, Tom  
; APPLICANT: Carter, Darrick  
; APPLICANT: Retter, Marc  
; APPLICANT: Mannion, Jane  
; APPLICANT: Fan, Liqun  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
; TITRATION OF INVENTION: DIAGNOSIS OF LUNG CANCER  
; FILE REFERENCE: 210121.478C14  
; CURRENT APPLICATION NUMBER: US/09/702,705  
; CURRENT FILING DATE: 2000-10-30  
; NUMBER OF SEQ ID NOS: 1833  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 966  
; LENGTH: 246  
; TYPE: DNA  
; ORGANISM: Homo sapien

US-09-702-705-966  
Alignment Scores:  
Pred. No.: 9.18e-68 Length: 246  
Score: 82.00 Matches: 82  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 5.74% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-702-705-966 (1-246)  
Qy 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProTyrValGly 858  
Db 246 CTGGGATAATGACGATCATCTCAGAGCATGTCAGGCTCACTTAACCCCATGGTGGC 187  
Qy 859 LeuArgLysIleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878  
Db 186 CTTCGGAAGATCAATGTCTCTACTGTGTCGTGGGAAAGATATGTCCTCCATTAAGT 127  
Qy 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898  
Db 126 TTATACAGTGGATGCGCTCTGAGCCCATGATGCTGTAATTTTATCAGAA 67

```
Qy 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCCAGTACTCGGCGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGCTAGTGTCTGT 7
Qy 919 GluArg 920
Db 6 GAAAGG 1

RESULT 8
US-09-736-457-966/c
; Sequence 966, Application US/09736457
; Patent No. 6509448
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Wang, Aijun
; APPLICANT: Fanger, Gary
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C15
; NUMBER OF SEQ ID NOS: 1864
; CURRENT FILING DATE: 2000-12-13
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-736-457-966

Alignment Scores:
Pred. No.: 9,18e-68 Length: 246
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.74% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-736-457-966 (1-246)
Qy 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrProTTPValGly 858
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCACCTTAACCCCACTCAATGCTAGTGTCTGT 187
Qy 859 LeuArgLysIleAsnValSerTyrTTPCysTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTTGCGAAGATCAATGTCTCTACTGTGCTGGAGCATGTCACCTTAACCCCACTCAATGCTAGTGTCTGT 127
Qy 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACAGTGGATGCCCTCTGAGCCAGTCAGTCGATGCTGTGGAATTTTATCAGAA 67
Qy 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCCAGTACTCGGCGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGCTAGTGTCTGT 7

US-09-787-097-12 (1-1429) x US-09-736-457-966 (1-246)
Qy 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrProTTPValGly 858
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCACCTTAACCCCACTCAATGCTAGTGTCTGT 187
Qy 859 LeuArgLysIleAsnValSerTyrTTPCysTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTTGCGAAGATCAATGTCTCTACTGTGCTGGAGCATGTCACCTTAACCCCACTCAATGCTAGTGTCTGT 127
Qy 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACAGTGGATGCCCTCTGAGCCAGTCAGTCGATGCTGTGGAATTTTATCAGAA 67
Qy 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCCAGTACTCGGCGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGCTAGTGTCTGT 7

US-09-736-457-966/c
; Sequence 966, Application US/09614124B
; Patent No. 6630574
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
```

```
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.478C9
; CURRENT APPLICATION NUMBER: US/09/614,124B
; NUMBER OF SEQ ID NOS: 1668
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-614-124B-966

Alignment Scores:
Pred. No.: 9,18e-68 Length: 246
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.74% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-614-124B-966 (1-246)
Qy 839 LeuArgIleMetGlnSerSerGlnSerMetSerLysLeuThrProTTPValGly 858
Db 246 CTGCGAATAATGCAGTCATCTCAGAGCATGTCACCTTAACCCCACTCAATGCTAGTGTCTGT 187
Qy 859 LeuArgLysIleAsnValSerTyrTTPCysTrpGluAspMetSerProPheThrAsnSer 878
Db 186 CTTGCGAAGATCAATGTCTCTACTGTGCTGGAGCATGTCACCTTAACCCCACTCAATGCTAGTGTCTGT 127
Qy 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlyIleLeuSerGlu 898
Db 126 TTACTACAGTGGATGCCCTCTGAGCCAGTCAGTCGATGCTGTGGAATTTTATCAGAA 67
Qy 899 ProSerThrArgGlyLeuLysAlaAlaThrCysIleAsnProLeuAsnGlySerValCys 918
Db 66 CCCAGTACTCGGCGACTGAAGGCTGCAACCTGCATCAACCCCACTCAATGCTAGTGTCTGT 7

US-09-614-124B-966/c
; Sequence 966, Application US/09671325
; Patent No. 6667154
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Lodes, Michael A.
; APPLICANT: Fanger, Gary
; APPLICANT: Vedvick, Tom
; APPLICANT: Carter, Darrick
; APPLICANT: Retter, Marc
; APPLICANT: Mannion, Jane
; APPLICANT: Fan, Liqun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.478C12
; CURRENT APPLICATION NUMBER: US/09/671,325
; NUMBER OF SEQ ID NOS: 1825
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 966
; LENGTH: 246
; TYPE: DNA
```

```
; ORGANISM: Homo sapien
US-09-671-325-966

Alignment Scores:
Pred. No.: 9,18e-68 Length: 246
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.74% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-671-325-966 (1-246)

QY 839 LeuArglleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProflpValGly 858
DB 246 CTGCGAATAATGAGTCTATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGC 187
QY 859 LeuArgLyslleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
DB 186 CTTCGGAGATCAATGTGTCTACTGTGTGTGGAGATATGCCCATTTTCAATAGT 127
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlylleLeuSerGlu 898
DB 126 TTACTACAGTGGATGCCGTCTGAGCCAGTGCATGCTGATTCTGTGGAATTTTATCAGAA 67
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCyslleAsnProLeuAsnGlySerValCys 918
DB 66 CCAGTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGTGTGTGTGT 7
QY 919 GluArg 920
DB 6 GAAAGG 1

RESULT 11
US-09-245-041-10
; Sequence 10, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 10
; LENGTH: 1051
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-10

Alignment Scores:
Pred. No.: 1,61e-61 Length: 1051
Score: 76.00 Matches: 1051
Percent Similarity: 98.23% Conservative: 0
Best Local Similarity: 98.23% Mismatches: 1
Query Match: 5.32% Indels: 2
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-245-041-10 (1-1051)

QY 87 AlaAlaValSerGlySerAlaAlaAlaGluAlaLysGluCysAspArgProCysVal 106
DB 463 GCGCGCGGTGTCCGCTCGCGCGAGCCAGGAGGATGTACCGCGCGGTGTGTC 522
QY 107 AsnGlyValArgCysAsnProGlyThrGlyGlnCysValCys-ProAlaGlyTrpValGlu 126

; ORGANISM: Homo sapien
US-09-671-325-966

Alignment Scores:
Pred. No.: 9,18e-68 Length: 246
Score: 82.00 Matches: 82
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.74% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-671-325-966 (1-246)

QY 839 LeuArglleMetGlnSerSerGlnSerMetSerLysLeuThrLeuThrProflpValGly 858
DB 246 CTGCGAATAATGAGTCTATCTCAGAGCATGTCCAAGCTCACCTTAACCCCATGGTGGC 187
QY 859 LeuArgLyslleAsnValSerTyrTrpCysTrpGluAspMetSerProPheThrAsnSer 878
DB 186 CTTCGGAGATCAATGTGTCTACTGTGTGTGGAGATATGCCCATTTTCAATAGT 127
QY 879 LeuLeuGlnTrpMetProSerGluProSerAspAlaGlyPheCysGlylleLeuSerGlu 898
DB 126 TTACTACAGTGGATGCCGTCTGAGCCAGTGCATGCTGATTCTGTGGAATTTTATCAGAA 67
QY 899 ProSerThrArgGlyLeuLysAlaAlaThrCyslleAsnProLeuAsnGlySerValCys 918
DB 66 CCAGTACTCGGGAGCTGAAGGCTGCAACCTGCATCAACCCACTCAATGTGTGTGTGTGT 7
QY 919 GluArg 920
DB 6 GAAAGG 1

RESULT 12
US-09-245-041-5
; Sequence 5, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; EARLIER FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 90050
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-5

Alignment Scores:
Pred. No.: 8,84e-57 Length: 90050
Score: 73.00 Matches: 73
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 5.11% Indels: 0
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-245-041-5 (1-90050)

QY 849 SerLysLeuThrLeuThrProTrpValGlyLeuArgLyslleAsnValSerTyrTrpCys 868
DB 28996 TCCAAGCTCACTGACTCCATGCGTGTGCTTCCGAGAGATCAATGTCTTACTGGTGC 29055
QY 869 TrpGluAspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSer 888
DB 29056 TGGGAGGATATGTCTCCATTCAAAATAGTTTGTGCGAGTGGATGCCATCTGAGCCCACT 29115
QY 889 AspAlaGlyPheCysGlylleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThr 908
DB 29116 GATGCTGGCTTCTGTGGGATCTTGTTCAGAGCTAGTACTCGGGGATTAAGGCTGCAACC 29175
QY 909 CyslleAsnProLeuAsnGlySerValCysGluArgPro 921
DB 29176 TGCATCAACCTCTCATATGGCAGCGCTCTGTGAAGGCTT 29214

RESULT 13
US-09-245-041-6/c
; Sequence 6, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
```

```

; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 6
; LENGTH: 226
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-6

Alignment Scores:
Pred. No.: 8,49e-39 Length: 226
Score: 51.00 Matches: 51
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.57% Indels: 0
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-245-041-6 (1-226)

QY 871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
Db 225 GATATGTCCTCCATTCAAAATAGTTTCTGTCAGTGGATGCACTGAGCCAGTGATGCT 166
QY 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
Db 165 GGCCTTCTGTGGATCTTGTTCAGAGCCTAGTACTCGGGGATTAAGGCTGAACCTGCATC 106
QY 911 AsnProLeuAnGlySerValCysGluArgPro 921
Db 105 AACCTCTCAATGGCAGCGCTGTGAAAGGCT 73

RESULT 14
US-09-245-041-7/c
; Sequence 7, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 221
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-7

Alignment Scores:
Pred. No.: 3,41e-33 Length: 221
Score: 45.00 Matches: 45
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 3.15% Indels: 0
DB: 3 Gaps: 0

```

```

US-09-787-097-12 (1-1429) x US-09-245-041-7 (1-221)

QY 871 AspMetSerProPheThrAsnSerLeuLeuGlnTrpMetProSerGluProSerAspAla 890
Db 220 GATATGTCCTCCATTCAAAATAGTTTCTGTCAGTGGATGCACTGAGCCAGTGATGCT 161
QY 891 GlyPheCysGlyIleLeuSerGluProSerThrArgGlyLeuLysAlaAlaThrCysIle 910
Db 160 GGCCTTCTGTGGATCTTGTTCAGAGCCTAGTACTCGGGGATTAAGGCTGAACCTGCATC 101
QY 911 AsnProLeuAnGly 915
Db 100 AACCTCTCAATGGT 86

RESULT 15
US-09-245-041-4
; Sequence 4, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 5973
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-4

Alignment Scores:
Pred. No.: 1,55e-20 Length: 5973
Score: 33.00 Matches: 33
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.31% Indels: 0
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-245-041-4 (1-5973)

QY 166 ProAsnArgIleMetArgLeuArgPheAsnHisPheAlaThrGluCysSerTrpAspHis 185
Db 4328 CCAAAATAGATAATAGAGACTTCGCTTCAACCATTTTGTACAGAAATGTAGCTGGACCAT 4387
QY 186 LeuTyValTyAspGlyAspSerIleTyAlaProLeu 198
Db 4388 TTATATGTTTATGATGGGAGCTCAATCTACGACACTCTG 4426

RESULT 16
US-09-245-041-3/c
; Sequence 3, Application US/09245041
; Patent No. 6274339
; GENERAL INFORMATION:
; APPLICANT: Moore, K.
; APPLICANT: Nagle, D.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE DIAGNOSIS AND TREATMENT
; TITLE OF INVENTION: OF BODY WEIGHT DISORDERS INCLUDING OBESITY
; FILE REFERENCE: 7853-136
; CURRENT APPLICATION NUMBER: US/09/245,041
; CURRENT FILING DATE: 1999-02-05
; EARLIER APPLICATION NUMBER: 60/093,630
; EARLIER FILING DATE: 1998-07-21
; EARLIER APPLICATION NUMBER: 60/104,978
; EARLIER FILING DATE: 1998-10-20
; NUMBER OF SEQ ID NOS: 131

```

```

; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 17056
; TYPE: DNA
; ORGANISM: Mus musculus
US-09-245-041-3

Alignment Scores:
Pred. No.: 2,84e-17 Length: 17056
Score: 30.00 Matches: 30
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 2.10% Indels: 0
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-245-041-3 (1-17056)

Qy 337 LeuLysLeuProArgAlaSerHisLysAlaValValAsnGlyAsnIleMetTTPValVal 356
Db 2473 TTAAGCTTCCAGAGCCTCTCATAAGCTGTGGTCAATGGAAATATAATGTGGTGTGT 2414

Qy 357 GlyGlyTyrMetPheAsnHisSerAspTyr 366
Db 2413 GCGGATATATGTTCAACCATTCAGATTAC 2384

RESULT 17
US-09-833-381-1916
; Sequence 1916, Application US/09833381
; Patent No. 6672186
; GENERAL INFORMATION:
; APPLICANT: Robison, Keith E.
; TITLE OF INVENTION: No. 6672186el Nucleic Acid and Protein Homologs
; FILE REFERENCE: 5800-119
; CURRENT APPLICATION NUMBER: US/09/833,381
; CURRENT FILING DATE: 2001-04-11
; PRIOR APPLICATION NUMBER: 09/516,448
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 2050
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1916
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)...(415)
; OTHER INFORMATION: n = A,T,C or G
US-09-833-381-1916

Alignment Scores:
Pred. No.: 2.11e-11 Length: 415
Score: 22.00 Matches: 22
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 1.54% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-833-381-1916 (1-415)

Qy 222 ThrSerGlyTyrAlaLeuLeuHisPheSerAspAlaAlaTyrAsnLeuThrGlyPhe 241
Db 247 ACATCTGGCTATGCACCTGTTACATTTTATTAGTGTGCTGCATTAATCTAAGTGGTTTC 306

Qy 242 AsnIle 243
Db 307 AACATT 312

RESULT 18
US-08-905-223-61
; Sequence 61, Application US/08905223
; Patent No. 6222029
; GENERAL INFORMATION:
; APPLICANT: Edwards, Jean-Baptiste D.
; TITLE OF INVENTION: N-PROTEINASE, AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
```

```

; APPLICANT: Duclert, Aymeric
; APPLICANT: Lacroix, Bruno
; TITLE OF INVENTION: 5' ESTs FOR SECRETED PROTEINS
; NUMBER OF SEQUENCES: 503
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson & Bear
; STREET: 501 West Broadway
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-3505
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Win95
; SOFTWARE: Word
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/905,223
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 61:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 429 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: DOUBLE
; TOPOLOGY: LINEAR
; MOLECULE TYPE: CDNA
; ORIGINAL SOURCE:
; ORGANISM: Homo Sapiens
; TISSUE TYPE: Brain
; FEATURE:
; NAME/KEY: sig_peptide
; LOCATION: 208..284
; IDENTIFICATION METHOD: Von Heijne matrix
; OTHER INFORMATION: score 7.6
; OTHER INFORMATION: seq LLNLLLSLFLAGL/DP
US-08-905-223-61

Alignment Scores:
Pred. No.: 0.00575 Length: 429
Score: 13.00 Matches: 13
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.91% Indels: 0
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-905-223-61 (1-429)

Qy 61 LeuLeuLeuLeuLeuLeuLeuProProLeuLeu 73
Db 54 CTGCTGTTATTGCTACTGCTGCTGCCGCCGCTCTGCTT 92

RESULT 19
US-09-491-522-6
; Sequence 6, Application US/09491522
; Patent No. 6428998
; GENERAL INFORMATION:
; APPLICANT: Colige, Alain
; APPLICANT: Lapiere, Charles M.
; APPLICANT: Prockop, Darwin J.
; TITLE OF INVENTION: RECOMBINANT N-PROTEINASE,
; AND THE PRODUCTION, METHODS AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
```





```

; TOPOLOGY: linear
;
; NAME/KEY: CDS
; FEATURE:
; Join(289..429, 433..477, 481..492, 496..603, 607
; LOCATION: .630, 634..750, 754..810, 814..843, 847..1023,
; LOCATION: 1027..1071, 1075..1103)
US-08-927-219-54
Alignment Scores:
Pred. No.: 0.127 Length: 1103
Score: 12.00 Matches: 12
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.84% Indels: 0
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-927-219-54 (1-1103)
QY 62 LeuLeuLeuLeuLeuLeuProProLeuLeu 73
Db 878 TTGTCCTCTCTCTGCTGCTGCCCACTCTGCTG 913

RESULT 23
US-09-028-327-2
; Sequence 2, Application US/09028327
; Patent No. 6130064
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl J.
; TITLE OF INVENTION: HUMAN SMN-LIKE PROTEIN
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/028,327
; FILING DATE:
; CLASSIFICATION:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0482 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2426 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRSTNOT24
; CLONE: 3769729
US-09-028-327-2
Alignment Scores:
Pred. No.: 0.28 Length: 2426
Score: 12.00 Matches: 12

```

Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.84% Indels: 0  
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-028-327-2 (1-2426)

QY 61 LeuLeuLeuLeuLeuLeuLeuProProLeu 72  
DB 507 CTGCTACTGCTACTGCTGCTGCTGCCACCGCCACTA 542

RESULT 24  
US-09-571-078A-2  
Sequence 2, Application US/09571078A  
Patent No. 6620783  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
Corley, Neil C.  
Guegler, Karl J.  
TITLE OF INVENTION: HUMAN SYN-LIKE PROTEIN  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Jacye Pharmaceuticals, Inc.  
STREET: 3174 Porter Dr.  
CITY: Palo Alto  
STATE: CA  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/571,078A  
FILING DATE: 15-May-2000  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Billings, Lucy J.  
REGISTRATION NUMBER: 36,749  
REFERENCE/DOCKET NUMBER: PP-0482 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 650-855-0555  
TELEFAX: 650-845-4166  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2426 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: BRSTNOT24  
CLONE: 3769729  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-571-078A-2

Alignment Scores:  
Pred. No.: 0.28 Length: 2426  
Score: 12.00 Matches: 12  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.84% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-571-078A-2 (1-2426)

QY 61 LeuLeuLeuLeuLeuLeuLeuProProLeu 72  
DB 507 CTGCTACTGCTACTGCTGCTGCTGCCACCGCCACTA 542

RESULT 25  
US-09-336-536-57  
Sequence 57, Application US/09336536  
Patent No. 6406884  
GENERAL INFORMATION:  
APPLICANT: Leiby, K.  
APPLICANT: McKay, C.  
APPLICANT: Bossone, S.  
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF  
FILE REFERENCE: 7853-144  
CURRENT APPLICATION NUMBER: US/09/336,536  
CURRENT FILING DATE: 1999-08-18  
NUMBER OF SEQ ID NOS: 75  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 57  
LENGTH: 639  
TYPE: DNA  
ORGANISM: Mus musculus  
US-09-336-536-57

Alignment Scores:  
Pred. No.: 0.636 Length: 639  
Score: 11.00 Matches: 11  
Percent Similarity: 100.00% Conservatives: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.77% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-336-536-57 (1-639)

QY 61 LeuLeuLeuLeuLeuLeuProPro 71  
DB 31 CTATTGCTGCTGCTGCTGCTGCTGCCGCCCT 63

RESULT 26  
US-09-620-312D-420/c  
Sequence 420, Application US/09620312D  
Patent No. 6569862  
GENERAL INFORMATION:  
APPLICANT: Tang, Y. Tom  
APPLICANT: Liu, Chenghua  
APPLICANT: Asundi, Vinod  
APPLICANT: Zhang, Jie  
APPLICANT: Ren, Feiyan  
APPLICANT: Chen, Rui-hong  
APPLICANT: Zhao, Qing A.  
APPLICANT: Wehrman, Tom  
APPLICANT: Xue, Aidong J.  
APPLICANT: Yang, Yonghong  
APPLICANT: Wang, Jian-Rui  
APPLICANT: Zhou, Ping  
APPLICANT: Ma, Yunqing  
APPLICANT: Wang, Dunrui  
APPLICANT: Wang, Zhiwei  
APPLICANT: John Tillinghast  
APPLICANT: Drmanac, Radoje T.  
TITLE OF INVENTION: No. 6569862el Nucleic Acids and  
FILE REFERENCE: 784CIP2B  
CURRENT APPLICATION NUMBER: US/09/620,312D  
CURRENT FILING DATE: 2000-07-19  
PRIOR APPLICATION NUMBER: 09/552,317  
PRIOR FILING DATE: 2000-04-25  
PRIOR APPLICATION NUMBER: 09/488,725  
PRIOR FILING DATE: 2000-01-21  
NUMBER OF SEQ ID NOS: 1105  
SOFTWARE: PC\_FL\_Genes Version 1.0  
SEQ ID NO 420  
LENGTH: 1432  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:

```
; NAME/KEY: CDS
; LOCATION: (59)..(1030)
US-09-620-312D-420

Alignment Scores:
Pred. No.: 1-43 Length: 1432
Score: 11.00 Matches: 11
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.77% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-620-312D-420 (1-1432)

Qy 64 LeuLeuLeuLeuProProProLeuLeuLeu 74
Db 628 CTCCTCTGCTGCTGCTGCTGCTGCTGCTC 596

RESULT 27
US-09-148-545-11
; Sequence 11, Application US/09148545
; Patent No. 6590075
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 70 Human Secreted Proteins
; FILE REFERENCE: P2001P1
; CURRENT APPLICATION NUMBER: US/09/148,545
; EARLIER FILING DATE: 1998-09-04
; EARLIER APPLICATION NUMBER: PCT/US98/04482
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,161
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587

; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
```



```
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; FILE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/907,794A
; CURRENT FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 194
; LENGTH: 2362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-794A-194

Alignment Scores:
Pred. No.:      2.35          Length:      2362
Score:          11.00        Matches:     11
Percent Similarity: 100.00% Conservative:    0
Best Local Similarity: 100.00% Mismatches:     0
Query Match:      0.77% Indels:         0
DB:               4 Gaps:         0

US-09-787-097-12 (1-1429) x US-09-907-794A-194 (1-2362)

Qy      61 LeuLeuLeuLeuLeuLeuProPro 71
       |||||TGTGCTGGTGCTGCTGCGCCGCCG 138
Db      106 CTGCTGTGCTGTGCTGCTGCTGCGCCGCCG 138

RESULT' 32
US-09-905-125A-194
; Sequence 194, Application US/09905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnovers, Luc
```



















TELEFAX: 509-927-1705  
TELEX:  
INFORMATION FOR SEQ ID NO: 132:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 100 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-08-811-492-132

Alignment Scores:  
Pred. No.: 0.859 Length: 100  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 10  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-811-492-132 (1-100)  
Qy 59 ProArgLeuLeuLeuLeuLeuLeu 68  
Db 42 CCGAGGTGTGTATTATTGTTGTTG 13

RESULT 54  
US-08-811-492-133/c  
Sequence 133, Application US/08811492  
Patent No. 5834247  
GENERAL INFORMATION:  
APPLICANT: COMB, DONALD G.  
APPLICANT: PERLER, FRANCES B.  
APPLICANT: JACK, WILLIAM E.  
APPLICANT: XU, MING-QUIN  
APPLICANT: HODGES, ROBERT A.  
APPLICANT: NOREN, CHRISTOPHER J.  
APPLICANT: CHONG, SHAO-RONG S.C.  
APPLICANT: ADAM, ERIC  
APPLICANT: SOUTHWORTH, MAURICE  
TITLE OF INVENTION: MODIFIED PROTEINS, METHODS OF THEIR  
TITLE OF INVENTION: PRODUCTION AND METHODS FOR PURIFICATION OF TARGET  
TITLE OF INVENTION: PROTEINS  
NUMBER OF SEQUENCES: 155  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: GREGORY D. WILLIAMS; NEW ENGLAND BIOLABS, INC.  
STREET: 32 TOZER ROAD  
CITY: BEVERLY  
STATE: MASSACHUSETTS  
COUNTRY: USA  
ZIP: 01915  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC\DOS\MS\DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/811,492  
FILING DATE: 28-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/580,555  
FILING DATE: 29-DEC-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/495,247  
FILING DATE: 28-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/146,885  
FILING DATE: 03-NOV-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/004,139

FILING DATE: 09-DEC-1992  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Williams, Gregory D.  
REGISTRATION NUMBER: 30901  
REFERENCE/DOCKET NUMBER: NEB-036C4  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 508-927-5054  
TELEFAX: 509-927-1705  
TELEX:  
INFORMATION FOR SEQ ID NO: 133:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 109 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: Genomic DNA  
US-08-811-492-133

Alignment Scores:  
Pred. No.: 0.936 Length: 109  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 2 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-811-492-133 (1-109)  
Qy 59 ProArgLeuLeuLeuLeuLeuLeu 68  
Db 81 CCGAGGTGTGTATTATTGTTGTTG 52

RESULT 55  
US-08-478-312-55/c  
Sequence 55, Application US/08478312  
Patent No. 5654276  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: England, Bruce  
APPLICANT: Schatz, Peter  
APPLICANT: Sloan, Derek  
APPLICANT: Chen, Min-Jia  
TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5  
TITLE OF INVENTION: Receptor  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies, N.V.  
STREET: 4001 Miranda Ave.  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/478,312  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691  
REFERENCE/DOCKET NUMBER: 1088.1A  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 55:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 122 base pairs  
TYPE: nucleic acid



STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (oligonucleotide)  
US-08-478-312-55

Alignment Scores:  
Pred. No.: 1.05 Length: 122  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 1 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-478-312-55 (1-122)

QY 59 ProArgLeuLeuLeuLeuLeuLeu 68  
DB 53 CCGAGTGTGTTATTATTATTATTGTTG 24

RESULT 56

US-08-485-302-55/c  
Sequence 55, Application US/08485302  
Patent No. 5668110  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: England, Bruce  
APPLICANT: Schatz, Peter  
APPLICANT: Sloan, Derek  
APPLICANT: Chen, Min-Jia  
TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5  
TITLE OF INVENTION: Receptor  
NUMBER OF SEQUENCES: 59  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies, N.V.  
STREET: 4001 Miranda Ave.  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,302  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691  
REFERENCE/DOCKET NUMBER: 1088.1B  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 55:  
LENGTH: 122 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (oligonucleotide)

US-08-485-302-55  
Alignment Scores:  
Pred. No.: 1.05 Length: 122  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 1 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-485-302-55 (1-122)

QY 59 ProArgLeuLeuLeuLeuLeuLeu 68  
DB 53 CCGAGTGTGTTATTATTATTATTGTTG 24

RESULT 57

US-08-476-169-60/c  
Sequence 60, Application US/08476169  
Patent No. 5677280  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: England, Bruce  
APPLICANT: Schatz, Peter  
APPLICANT: Sloan, Derek  
APPLICANT: Chen, Min-Jia  
TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5  
TITLE OF INVENTION: Receptor  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies, N.V.  
STREET: 4001 Miranda Ave.  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/476,169  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691  
REFERENCE/DOCKET NUMBER: 1088.2A  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 122 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (oligonucleotide)

US-08-476-169-60  
Alignment Scores:  
Pred. No.: 1.05 Length: 122  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 1 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-476-169-60 (1-122)

QY 59 ProArgLeuLeuLeuLeuLeuLeu 68  
DB 53 CCGAGTGTGTTATTATTATTATTGTTG 24

RESULT 58

US-08-484-083-60/c  
Sequence 60, Application US/08484083  
Patent No. 5683983  
GENERAL INFORMATION:  
APPLICANT: Barrett, Ronald W.  
APPLICANT: England, Bruce  
APPLICANT: Schatz, Peter  
APPLICANT: Sloan, Derek

APPLICANT: Chen, Min-Jia  
TITLE OF INVENTION: Peptides and Compounds That Bind to the IL-5  
TITLE OF INVENTION: Receptor  
NUMBER OF SEQUENCES: 65  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Affymax Technologies, N.V.  
STREET: 4001 Miranda Ave.  
CITY: Palo Alto  
STATE: California  
COUNTRY: USA  
ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/484,083  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Stevens, Lauren L.  
REGISTRATION NUMBER: 36,691  
REFERENCE/DOCKET NUMBER: 1088.2B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-496-2300  
TELEFAX: 415-424-0832  
INFORMATION FOR SEQ ID NO: 60:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 122 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (oligonucleotide)  
US-08-484-083-60  
Alignment Scores:  
Pred. No.: 1.05 Length: 122  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 1 Gaps: 0  
US-09-787-097-12 (1-1429) x US-08-484-083-60 (1-122)  
QY 59 ProArgLeuLeuLeuLeuLeuLeu 58  
Db 53 CCGAGGTGTGTATTGTTATTGTTG 24  
RESULT 59  
US-08-764-640-237/c  
Sequence 237, Application US/08764640  
Patent No. 5869451  
Patent No. 5869451 5837693  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwila, Steven E.  
APPLICANT: Gates, Christian  
APPLICANT: Schatz, Peter J.  
APPLICANT: Balasubramanian, Palaniappan  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Hendren, Richard W.  
APPLICANT: Depirnce, Randolph B.  
APPLICANT: Podduturi, Surekha  
APPLICANT: Yin, Qun  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
TITLE OF INVENTION: RECEPTOR  
NUMBER OF SEQUENCES: 244  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398

CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/764,640  
FILING DATE: 11-DEC-1996  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 237:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 122 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..120  
US-08-764-640-237  
Alignment Scores:  
Pred. No.: 1.05 Length: 122  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 2 Gaps: 0  
US-09-787-097-12 (1-1429) x US-08-764-640-237 (1-122)  
QY 59 ProArgLeuLeuLeuLeuLeuLeu 68  
Db 53 CCGAGGTGTGTATTGTTATTGTTG 24  
RESULT 60  
US-08-973-225-225/c  
Sequence 225, Application US/08973225A  
Patent No. 6083913  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwila, Steven E.  
APPLICANT: Duffin, David J.  
APPLICANT: Gates, Christian  
APPLICANT: Haselden, Sherril S.  
APPLICANT: Mattheakis, Larry C.  
APPLICANT: Schatz, Peter J.  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Wrighton, Nicholas C.  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
TITLE OF INVENTION: THROMBOPOIETIN RECEPTOR  
NUMBER OF SEQUENCES: 232  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/973,225A  
FILING DATE: 04-Dec-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3065USW  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 225:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 122 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..120  
SEQUENCE DESCRIPTION: SEQ ID NO: 225:  
US-08-973-225-225  
Alignment Scores:  
Pred. No.: 1.05 Length: 122  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 3 Gaps: 0  
US-09-787-097-12 (1-1429) x US-08-973-225-225 (1-122)  
QY 59 ProArgLeuLeuLeuLeuLeuLeu 68  
DB 53 CCGAGTGTGTTGTTATTGTTGTTG 24  
RESULT 61  
US-09-244-298A-237/c  
Sequence 237, Application US/09244298A  
Patent No. 6121238  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwiria, Steven E.  
APPLICANT: Gates, Christian  
APPLICANT: Schatz, Peter J.  
APPLICANT: Balasubramanian, Palaniappan  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Hendren, Richard W.  
APPLICANT: Deprince, Randolph B.  
APPLICANT: Poduturi, Surekha  
APPLICANT: Yin, Qun  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
TITLE OF INVENTION: RECEPTOR  
NUMBER OF SEQUENCES: 244  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/244,298A  
FILING DATE: 11-DEC-1996  
CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 237:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 122 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..120  
US-09-244-298A-237  
Alignment Scores:  
Pred. No.: 1.05 Length: 122  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 3 Gaps: 0  
US-09-787-097-12 (1-1429) x US-09-244-298A-237 (1-122)  
QY 59 ProArgLeuLeuLeuLeuLeuLeu 68  
DB 53 CCGAGTGTGTTGTTATTGTTGTTG 24  
RESULT 62  
US-09-516-704-237/c  
Sequence 237, Application US/09516704  
Patent No. 6251864  
GENERAL INFORMATION:  
APPLICANT: Dower, William J.  
APPLICANT: Barrett, Ronald W.  
APPLICANT: Cwiria, Steven E.  
APPLICANT: Gates, Christian  
APPLICANT: Schatz, Peter J.  
APPLICANT: Balasubramanian, Palaniappan  
APPLICANT: Wagstrom, Christopher R.  
APPLICANT: Hendren, Richard W.  
APPLICANT: Deprince, Randolph B.  
APPLICANT: Poduturi, Surekha  
TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A  
TITLE OF INVENTION: RECEPTOR  
NUMBER OF SEQUENCES: 244  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Glaxo Wellcome  
STREET: Five Moore Drive, P.O. Box 13398  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/516,704  
FILING DATE: 01-Mar-2000  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Hrubiec, Robert T.  
REGISTRATION NUMBER: 36,392  
REFERENCE/DOCKET NUMBER: PK3281  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-248-1000  
INFORMATION FOR SEQ ID NO: 237:

```
;
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..120
; SEQUENCE DESCRIPTION: SEQ ID NO: 225:
US-09-516-704-237

Alignment Scores:
Pred. No.: 1.05 Length: 122
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.70% Indels: 0
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-516-704-237 (1-122)
QY 59 ProArgLeuLeuLeuLeuLeuLeu 68
DB 53 CCGAGGTGTGTTATTGTTGTTG 24

RESULT 63
US-09-549-090-225/c
; Sequence 225, Application US/09549090
; Patent No. 6465430
; GENERAL INFORMATION:
; APPLICANT: Dower, William J.
; Barrett, Ronald W.
; Cwirla, Steven E.
; Duffin, David J.
; Gates, Christian
; Haselden, Sherril S.
; Mattheakis, Larry C.
; Schatz, Peter J.
; Wagstrom, Christopher R.
; Wrighton, Nicholas C.
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
; THROMBOPOIETIN RECEPTOR
; NUMBER OF SEQUENCES: 232
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/549,090
; FILING DATE: 13-Apr-2000
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/973,225
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3065USW
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 225:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..120
; SEQUENCE DESCRIPTION: SEQ ID NO: 225:
US-09-832-230A-237

Alignment Scores:
Pred. No.: 1.05 Length: 122
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.70% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-832-230A-237 (1-122)
QY 59 ProArgLeuLeuLeuLeuLeuLeu 68
DB 53 CCGAGGTGTGTTATTGTTGTTG 24

RESULT 64
US-09-832-230A-237/c
; Sequence 237, Application US/09832230A
; Patent No. 6506362
; GENERAL INFORMATION:
; APPLICANT: Dower, William J. et al
; TITLE OF INVENTION: PEPTIDES AND COMPOUNDS THAT BIND TO A
; RECEPTOR
; NUMBER OF SEQUENCES: 244
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Glaxo Wellcome
; STREET: Five Moore Drive, P.O. Box 13398
; CITY: Research Triangle Park
; STATE: NC
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/832,230A
; FILING DATE: 10-Apr-2001
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Hrubiec, Robert T.
; REGISTRATION NUMBER: 36,392
; REFERENCE/DOCKET NUMBER: PK3281
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-248-1000
; INFORMATION FOR SEQ ID NO: 237:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 122 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..120
; SEQUENCE DESCRIPTION: SEQ ID NO: 237:
US-09-832-230A-237

Alignment Scores:
Pred. No.: 1.05 Length: 122
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.70% Indels: 0
DB: 4 Gaps: 0
```

US-09-787-097-12 (1-1429) x US-09-832-230A-237 (1-122)

QY 59 ProArgLeuLeuLeuLeuLeuLeuLeu 68  
Db 53 CCGAGGTTGTTGTTATGTTATGTTGTTG 24

## RESULT 65

US-08-938-548B-21  
; Sequence 21, Application US/08938548B  
; Patent No. 6001963  
; GENERAL INFORMATION:  
; APPLICANT: Yanagisawa, Masashi  
; APPLICANT: Bergsma, Derk  
; APPLICANT: Wilson, Shelagh  
; APPLICANT: Brooks, David  
; APPLICANT: Gellai, Miklos  
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE  
; TITLE OF INVENTION: RECEPTOR HFGAN72  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: United States of America  
; ZIP: 19406

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,548B  
FILING DATE: 26-SEPT-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/887,382  
FILING DATE: 2-JUL-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/820,519  
FILING DATE: 19-MAR-1997  
APPLICATION NUMBER: 60/033,604  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth J. Hecht  
REGISTRATION NUMBER: 41,824  
REFERENCE/DOCKET NUMBER: ATG50037-2  
TELEPHONE: 610-270-5009  
TELEFAX: 610-270-5090  
TELEX:

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 577 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-938-548B-21

Alignment Scores:  
Pred. No.: 4.95 Length: 577  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-938-548B-21 (1-577)

QY 61 LeuLeuLeuLeuLeuLeuLeuProPro 70  
Db |||||

US-08-938-548B-21  
; Sequence 21, Application US/08938548B  
; Patent No. 6001963  
; GENERAL INFORMATION:  
; APPLICANT: Yanagisawa, Masashi  
; APPLICANT: Bergsma, Derk  
; APPLICANT: Wilson, Shelagh  
; APPLICANT: Brooks, David  
; APPLICANT: Gellai, Miklos  
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE  
; TITLE OF INVENTION: RECEPTOR HFGAN72  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: United States of America  
; ZIP: 19406

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/938,548B  
FILING DATE: 26-SEPT-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/887,382  
FILING DATE: 2-JUL-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/820,519  
FILING DATE: 19-MAR-1997  
APPLICATION NUMBER: 60/033,604  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Elizabeth J. Hecht  
REGISTRATION NUMBER: 41,824  
REFERENCE/DOCKET NUMBER: ATG50037-2  
TELEPHONE: 610-270-5009  
TELEFAX: 610-270-5090  
TELEX:

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 577 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-938-548B-21

Alignment Scores:  
Pred. No.: 4.95 Length: 577  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 3 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-938-548B-21 (1-577)

QY 61 LeuLeuLeuLeuLeuLeuLeuProPro 70  
Db |||||

Db 130 CTACTGCTGCTGCTGCTGCTGCTGCCGCC 159

## RESULT 66

US-08-939-093A-21  
; Sequence 21, Application US/08939093A  
; Patent No. 6309854  
; GENERAL INFORMATION:  
; APPLICANT: Yanagisawa, Masashi  
; APPLICANT: Bergsma, Derk  
; APPLICANT: Wilson, Shelagh  
; APPLICANT: Brooks, David  
; APPLICANT: Gellai, Miklos  
; TITLE OF INVENTION: NOVEL LIGANDS OF THE NEUROPEPTIDE  
; TITLE OF INVENTION: RECEPTOR HFGAN72  
; NUMBER OF SEQUENCES: 21  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: SmithKline Beecham Corporation  
; STREET: 709 Swedeland Road  
; CITY: King of Prussia  
; STATE: PA  
; COUNTRY: United States of America  
; ZIP: 19406

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/939,093A  
FILING DATE: 26-SEPT-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/887,382  
FILING DATE: 2-JUL-1997

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/820,519  
FILING DATE: 19-MAR-1997  
APPLICATION NUMBER: 60/033,604  
FILING DATE: 17-DEC-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: King, William T.  
REGISTRATION NUMBER: 30,954  
REFERENCE/DOCKET NUMBER: ATG50037-3  
TELEPHONE: 610-270-5219  
TELEFAX: 610-270-4026  
TELEX:

INFORMATION FOR SEQ ID NO: 21:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 577 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-939-093A-21

Alignment Scores:  
Pred. No.: 4.95 Length: 577  
Score: 10.00 Matches: 10  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 0.70% Indels: 0  
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-939-093A-21 (1-577)

QY 61 LeuLeuLeuLeuLeuLeuLeuProPro 70  
Db |||||

US-09-787-097-12 (1-1429) x US-08-939-093A-21 (1-577)

QY 61 LeuLeuLeuLeuLeuLeuLeuProPro 70

Db 130 CTACTGCTGCTGCTGCTGCTGCCGCC 159

RESULT 67

US-09-211-823C-21

; Sequence 21, Application US/09211823C



```
;
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-811-481-26

Alignment Scores:
Pred. No.: 8.57 Length: 999
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.70% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-08-811-481-26 (1-999)

QY 61 LeuLeuLeuLeuLeuLeuProPro 70
DB 1 CTGTTGCTGCTACTGCTGCTGCTGCCGCCA 30

RESULT 71
US-09-787-097-12 (1-1429) x US-08-811-481-26 (1-999)
; Sequence 26, Application US/09876527
; Patent No. 6627735
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; Jelinek, Laura J.
; Sheppard, Paul O.
; Hagopian, William A.
; LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/876,527
; FILING DATE: 07-Jun-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/811,481
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 999 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
```

```
;
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 26:
US-09-876-527-26

Alignment Scores:
Pred. No.: 8.57 Length: 999
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.70% Indels: 0
DB: 4 Gaps: 0

US-09-787-097-12 (1-1429) x US-09-876-527-26 (1-999)

QY 61 LeuLeuLeuLeuLeuLeuProPro 70
DB 1 CTGTTGCTGCTACTGCTGCTGCTGCCGCCA 30

RESULT 72
US-08-811-481-27
; Sequence 27, Application US/08811481
; Patent No. 6300093
; GENERAL INFORMATION:
; APPLICANT: Kindsvogel, Wayne
; Jelinek, Laura J.
; Sheppard, Paul O.
; Hagopian, William A.
; LaGasse, James M.
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/811,481
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lingenfelter, Susan
; REGISTRATION NUMBER: P-41,156
; REFERENCE/DOCKET NUMBER: 95-36
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-442-6675
; TELEFAX: 206-442-6678
; TELEX:
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1011 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; US-08-811-481-27

Alignment Scores:
Pred. No.: 8.68 Length: 1011
Score: 10.00 Matches: 10
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 0.70% Indels: 0
```



DB: 4 Gaps: 0  
US-09-787-097-12 (1-1429) x US-08-811-481-27 (1-1011)  
Qy 61 LeuLeuLeuLeuLeuLeuLeuLeuProPro 70  
Db 13 CTGTTGCTGCTACTGCTGCTGCTGCCGCCA 42  
RESULT 73  
US-09-876-527-27  
; Sequence 27, Application US/09876527  
; Patent No. 6627735  
; GENERAL INFORMATION:  
; APPLICANT: Kindsvogel, Wayne  
; Jelinek, Laura J.  
; Sheppard, Paul O.  
; Hagopian, William A.  
; LaGasse, James M.  
; TITLE OF INVENTION: ISLET CELL ANTIGEN 1851  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 1201 Eastlake Avenue East  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/876,527  
; FILING DATE: 07-Jun-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/811,481  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Lingenfelter, Susan  
; REGISTRATION NUMBER: P-41,156  
; REFERENCE/DOCKET NUMBER: 95-36  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 206-442-6675  
; TELEFAX: 206-442-6678  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1011 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 27:  
US-09-876-527-27  
Alignment Scores: 8.68 Length: 1011  
Pred. No.: 10.00 Matches: 10  
Score: 100.00%  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 0.70%  
Indels: 0  
Gaps: 0  
US-09-787-097-12 (1-1429) x US-09-876-527-27 (1-1011)  
Qy 61 LeuLeuLeuLeuLeuLeuLeuLeuProPro 70  
Db 13 CTGTTGCTGCTACTGCTGCTGCTGCCGCCA 42  
RESULT 74  
US-09-489-039A-3636/c

; Sequence 3636, Application US/09489039A  
; Patent No. 6810836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Bretton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709,2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; PRIOR FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 3636  
; LENGTH: 1146  
; TYPE: DNA  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-3636  
Alignment Scores: 9.84 Length: 1146  
Pred. No.: 10.00 Matches: 10  
Score: 100.00%  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 0.70%  
Indels: 0  
Gaps: 0  
US-09-787-097-12 (1-1429) x US-09-489-039A-3636 (1-1146)  
Qy 89 AlaValSerGlySerAlaAlaAlaGluAla 98  
Db 521 GCCGTACGCGCAGCGCGCGCGGAGGCA 492  
RESULT 75  
US-09-177-325-2/c  
; Sequence 2, Application US/09177325B  
; Patent No. 6214983  
; GENERAL INFORMATION:  
; APPLICANT: Tureci, Ozlem  
; APPLICANT: Sakin, Ugur  
; TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, A  
; FILE REFERENCE: Thereof  
; FILE REFERENCE: LJD 5525  
; CURRENT APPLICATION NUMBER: US/09/177,325B  
; CURRENT FILING DATE: 1998-10-22  
; NUMBER OF SEQ ID NOS: 15  
; SEQ ID NO 2  
; LENGTH: 1276  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
US-09-177-325-2  
Alignment Scores: 11 Length: 1276  
Pred. No.: 10.00 Matches: 10  
Score: 100.00%  
Percent Similarity: 100.00%  
Best Local Similarity: 100.00%  
Query Match: 0.70%  
Indels: 0  
Gaps: 0  
US-09-787-097-12 (1-1429) x US-09-177-325-2 (1-1276)  
Qy 61 LeuLeuLeuLeuLeuLeuLeuProPro 70  
Db 813 CTCCTTCTCTCTCTCTCTCTCTCTCTCTCT 784  
Search completed: March 1, 2004, 13:06:12  
Job time : 751 secs